





CC (especially 5-100) mg/kg body weight. The humanised antibody has low antigenicity and is therefore effective therapeutically in humans.  
 CC XX  
 SQ Sequence 180 AA;

Query Match 100.0%; Score 889; DB 19; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-78;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MASTSYDCRVPMEGDGDKRCKLGLGIGILVLLILVILGVLIFTIKANSEACRDGLRAV 60  
 Db 1 mastsydcrvpmedgdkrckllglgigilvllilvlgpliftikanseacrdgirav 60  
 Qy 1 MASTSYDCRVPMEGDGDKRCKLGLGIGILVLLILVILGVLIFTIKANSEACRDGLRAV 60  
 Db 1 mastsydcrvpmedgdkrckllglgigilvllilvlgpliftikanseacrdgirav 60  
 Qy 61 MECRNVTHLQQELTEAQKGFDQVEAQATCNHTYMAALASLDKAQOKKVLEGEI 120  
 Qy 61 mecnvthlqqelteaqkgfdqveaqatcnhtymalasldkaqokkvlegei 120  
 Db 61 mecnvthlqqelteaqkgfdqveaqatcnhtymalasldkaqokkvlegei 120  
 Qy 121 TTLNHKLQDASAVERLRENQVLSVRIADKKYPPSSQDSSAAPOLIJVLGLSALLQ 180  
 Db 121 ttlnhkldqdasaverlrenqvisvriadkkypssqssaaapqljvlglsallq 180

RESULT 4  
 AAY33202 standard; Protein; 180 AA.  
 AC AAY33202;  
 XX  
 DT 22-NOV-1999 (first entry)  
 XX Human HML.24 antigenic protein.  
 DE XX  
 KW Antigenic Protein; HML.24; splice variant; promoter; antirheumatic;  
 KW antiarthritic; bone marrow; tumour cell; drug development; treatment;  
 KW myeloma; rheumatoid arthritis; human.  
 OS XX  
 PN Homo sapiens.  
 XX WO9943803-A1.  
 XX PD 02-SEP-1999.  
 XX PF 25-FEB-1999; 99WO-JP00884.  
 XX PR 25-FEB-1998; 98JP-0060617.  
 XX PR 24-MAR-1998; 98JP-0093883.  
 XX PA (CHUS ) CHUGAI SEIYAKU KK.  
 XX PI Koishihara Y, Ozaki Y;  
 XX PN WO9943703-A1.  
 XX DR 02-SEP-1999.  
 XX PR 25-FEB-1999; 99WO-JP00885.  
 XX PR 25-FEB-1998; 98JP-0060613.  
 XX PA (CHUS ) CHUGAI SEIYAKU KK.  
 XX PT Immunoassay of anti-HML.24 antibody or soluble HML.24 antigen,  
 XX PT useful for diagnosis of immune disorders and cancer  
 XX PS Example 3; Fig 14-15; 138pp; Japanese.  
 XX  
 CC This sequence represents a human soluble HML.24 antigenic protein. The  
 CC invention relates to an immunochemical assay of anti-HML.24 antibody by  
 CC use of a soluble HML.24 antigenic protein, or an immunochemical assay of  
 CC the soluble antigen by use of the antibody. The immunoassay of the HML.24  
 CC antigen or antibody is useful for diagnosis of immune disorders and  
 CC cancer, for monitoring of anti-HML.24 antibody immunotherapy, and for  
 CC assay of the antibody or antigen for investigative purposes, in  
 CC biological samples such as blood, serum, urine, milk, synovial fluid or  
 CC microorganism culture media. The method is sensitive down to 500 pg/ml  
 CC antibody.  
 XX Sequence 180 AA;

PT Genomic DNA encoding HML.24 antigen protein as well as splicing  
 PT variants, useful e.g. in development of drugs for treating myeloma and  
 PT rheumatoid arthritis  
 XX Example 1; Fig 1-2; 83pp; Japanese.  
 XX  
 CC This invention describes a novel human antigenic protein, HML.24,  
 CC its encoding nucleic acid, splice variants and promoter region. The  
 CC products of the invention have antirheumatic and antiarthritic activity.  
 CC The DNA of the invention is isolated from bone marrow tumour cells,  
 CC which can be used to study the expression of HML.24 antigen, promoter  
 CC activity of its promoter region, and in development of drugs in treating  
 CC e.g. myeloma and rheumatoid arthritis. This sequence represents the  
 CC human HML.24 antigenic protein described in the invention.  
 XX Sequence 180 AA;

Qy 1 MASTSYDCRVPMEGDGDKRCKLGLGIGILVLLILVILGVLIFTIKANSEACRDGLRAV 60  
 Db 1 mastsydcrvpmedgdkrckllglgigilvllilvlgpliftikanseacrdgirav 60

**RESULT 7**  
AY05484  
ID AY05484 standard; Protein; 180 AA.  
XX  
AC AY05484;  
XX  
DT 07-JUL-1999 (first entry)  
DE Potentiator for antibody against lymphoid tumour.  
XX  
KW Antibody potentiator; lymphoid tumour; lymphoma; cytotoxic antibody;  
KW multiple myeloma; acute B-lymphocytoma; chronic B-lymphocytoma;  
KW pre-B lymphoma; Burkitt's lymphoma; acute T-lymphocytoma; therapy;  
KW chronic T-lymphocytoma; PNTL.  
XX  
OS Homo sapiens.  
XX  
PN WO9918997-A1.  
XX  
PD 22-APR-1999.  
XX  
PF 14-OCT-1998; 98WO-JP04645.  
XX  
PR 05-AUG-1998; 98JP-0222024.  
PR 14-OCT-1997; 97JP-0280759.  
XX  
PA (CHUS ) CHUGAI SEIYAKU KK.  
XX  
PT Koishihara Y, Kosaka M;  
XX  
DR WO1; 1999-277447/23.  
XX  
DR N-PSDB; AAX36561.  
XX  
PT Potentiation of antibody treatment of lymphoma with biological  
PT response modifier.  
XX  
PS Claim 1; Page 37-38; 62pp; Japanese.  
XX  
CC This sequence represents a potentiator for an antibody against  
CC lymphoid tumour.  
CC The invention relates to a method for the treatment of lymphoma, in which  
CC a cytotoxic antibody is potentiated by administration of a biological  
CC response modifier. The method can be used for treatment of lymphomas and  
CC multiple myelomas which are resistant to conventional treatment, such as  
CC acute B-lymphocytoma, chronic B-lymphocytoma, pre-B lymphoma, Burkitt's  
CC lymphoma, acute T-lymphocytoma, chronic T-lymphocytoma, and PNTL.  
XX  
SQ Sequence 180 AA;  
  
Query Match 100.0%; Score 889; DB 20; Length 180;  
Best Local Similarity 100.0%; Pred. No. 6-2e-78; Indels 0;  
Matches 180; Conservative 0; Mismatches 0; Gaps 0;  
  
**QY** 1 MASTSYDCRYVMEDEGDKRCKULLGIGIILVLLILIVLGWLPLIFTIKANSEACRDGLRVA 60  
**Db** 1 mastsydcryvmedegdkrckullgigilvllilivlgwlpliftikansecdrglrv 60  
**QY** 61 MECRNVTHILLQBLTEAQKGFQDVEAQATCNCIHMALMASLDAEKAQGQKVEEGETI 120  
**Db** 61 mecrnvthillqblteaqkgfqdveaqatcnhtvmaimasiidaekaqgqkveegei 120  
**QY** 121 TPLNKHQDASAEVERIRRENQYLSVRIADKKYPPSSQDSSAAPOLILVIGLSALQ 180  
**Db** 121 tlnhkqdasaeverirrenqylsvriadkkypssqdsaaapqlilvlglsalq 180

AC	AY07250;	PF	20-AUG-1999;	99W00-JP04502.	
XX	06-JUL-1999 (first entry)	PR	18-SEP-1998;	98JP-0264593.	
DT		XX			
XX	BST-2 protein.	PA			
DE		XX			
XX	Mouse; BST-2; monoclonal antibody; RS38; myeloma; cytotoxic activity.	PI			
KW		XX			
XX	OS	DR			
XX	Mus sp.	DR			
PN	JP11092399-A.	N_PSDB; AAA13654.			
PD	06-APR-1999.				
PF	24-SEP-1997;	PT			
XX	97JP-0274960.	PT			
PR	24-SEP-1997;	XX			
XX	97JP-0274960.	PS			
PA	(CHUS ) CHUGAI PHARM CO LTD.	XX			
XX	WPI: 1999-283503/24.	CC			
DR	N_PSDB; AAX29996.	CC			
XX		CC			
PT	An agent for treating myeloma - includes an antibody and has cytotoxic activity	CC			
CC	This sequence represents the mouse BST-2 protein which is used to raise antibodies, especially the monoclonal antibody RS38. The antibody can be used in compositions to treat myelomas when the antibody is associated with a cytotoxic activity.	CC			
CC		CC			
PS	Claim 1; Page 10; 13pp; Japanese.	CC			
XX		CC			
CC		CC			
CC		CC			
XX	Sequence 180 AA;	CC			
Query	Match	100.0%	Score	889; DB 20; Length 180;	
Best	Local Similarity	100.0%	Pred.	No. 6.2e-78;	
Matches	180; Conservative	0; Mismatches	0; Indels	0; Gaps	
Qy	1 MASTSYDYCRVPMEGDCKRCKLIGIGLVLILITWIGVPLIFTIKANSEACRDGLRAV 60	Qy	1 MASTSYDYCRVPMEGDCKRCKLIGIGLVLILITWIGVPLIFTIKANSEACRDGLRAV 60	Qy	1 MASTSYDYCRVPMEGDCKRCKLIGIGLVLILITWIGVPLIFTIKANSEACRDGLRAV 60
Db	1 mastsydycrvpmedgdkrckliligglvllvlgvpiliftikansacrdglra 60	Db	1 mastsydycrvpmedgdkrckliligglvllvlgvpiliftikansacrdglra 60	Db	1 mastsydycrvpmedgdkrckliligglvllvlgvpiliftikansacrdglra 60
Qy	61 MECRNWTHLLQQLTEAQKGFDQVEAQATCNHTWMLMASLDAEKAOQKVKVELEGEI 120	Qy	61 MECRNWTHLLQQLTEAQKGFDQVEAQATCNHTWMLMASLDAEKAOQKVKVELEGEI 120	Qy	61 MECRNWTHLLQQLTEAQKGFDQVEAQATCNHTWMLMASLDAEKAOQKVKVELEGEI 120
Db	61 mecrnwthllqqlteaqkgfdqveaqatcnhtwmlmasladaekaqgkvvlegei 120	Db	61 mecrnwthllqqlteaqkgfdqveaqatcnhtwmlmasladaekaqgkvvlegei 120	Db	61 mecrnwthllqqlteaqkgfdqveaqatcnhtwmlmasladaekaqgkvvlegei 120
Qy	121 TTLNHKLQDASAEVERLRENQVLSVRAKDKYVSSQDSSAAPQLITVLLGSLALQ 180	Qy	121 TTLNHKLQDASAEVERLRENQVLSVRAKDKYVSSQDSSAAPQLITVLLGSLALQ 180	Qy	121 TTLNHKLQDASAEVERLRENQVLSVRAKDKYVSSQDSSAAPQLITVLLGSLALQ 180
Db	121 ttlnhklqdasaeverlrenqvlsvriakkypssqsdssaaapqlivlglisal 180	Db	121 ttlnhklqdasaeverlrenqvlsvriakkypssqsdssaaapqlivlglisal 180	Db	121 ttlnhklqdasaeverlrenqvlsvriakkypssqsdssaaapqlivlglisal 180
RESULT	10				
ID	ABP50295				
ID	ABP50295 standard; Protein; 180 AA.				
XX					
AC	ABP50295;				
XX					
DT	08-FEB-2002 (first entry)				
XX					
DE	Bone marrow stromal antigen (BST-2) ovarian tumour marker protein, #80.				
XX					
XX	Ovarian tumour marker gene; human; overexpression; upregulation; epithelial tumour; cancer; diagnosis; prognosis; disease monitoring; identification; serious cystadenocarcinoma; borderline serious tumour; serious cystadenocarcinoma; mucinous cystadenocarcinoma; mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma; undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma; adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE; immune response pathway; cell proliferation regulation; protein folding; membrane localised; secreted; therapeutic target; cytostatic; gene therapy; vaccine.				
KW	Homo sapiens.				
XX					
PN	WO200017395-A1.				
PD	30-MAR-2000.				
XX					

PN WO200175177-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PR 03-APR-2001; 2001WO-US10947.  
 XX  
 PR 03-APR-2000; 2000US-194336P.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;  
 XX  
 DR WPI; 2001-626450/72.  
 DR N-PSDB; ABA83121.  
 XX  
 PT Detecting and identifying ovarian tumor, identifying increased risk for developing ovarian cancer, and determining effectiveness of ovarian cancer treatment, by measuring expression level of ovarian tumor marker gene -  
 PT  
 XX  
 PS Claim 23; Page 124; 140pp; English.  
 XX  
 CC The invention relates to methods for diagnosing and prognosis ovarian tumours in an individual via the detection and measurement of the expression of ovarian tumour marker genes (ABA83081-ABA83122, ABA83180, ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179, ABA83181 and ABA83183). The methods of the invention are useful for detecting an ovarian tumour in a patient, for identifying an individual at increased risk for developing ovarian cancer, for prognostic tests for assessing the relative severity of ovarian cancer, in tests for monitoring a patient in remission from ovarian cancer and in tests for monitoring disease status in a patient being treated for ovarian cancer. The methods can additionally be used to identify a particular tumour as being an ovarian tumour (i.e., an epithelial ovarian tumour selected from serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma, mucinous cystadenoma, borderline mucinous tumour, mucinous cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma, clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner tumour. The ovarian tumour marker genes of the invention were identified using SAGE (serial analysis of gene expression) and were found to be overexpressed in a broad variety of ovarian epithelial tumour cells relative to normal ovarian epithelial cells. The marker genes are implicated in immune response pathways, in the regulation of cell proliferation and in protein folding, and many of these are membrane-localised or secreted. In addition to their use as diagnostic and prognostic markers, the ovarian tumour marker genes or their encoded proteins may be used as therapeutic targets for the treatment and prevention of ovarian cancer. Sequences ABB50257-ABB50299 represent CC proteins encoded by ovarian tumour marker genes of the invention. CC  
 XX  
 SQ Sequence 180 AA;

Query Match 100.0%; Score 889; DB 22; Length 180; Best Local Similarity 100.0%; Pred. 6.2e-78; Indels 0; Gaps 0; Matches 180; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 MASTSVDICRYPMEDGDKRCKLLGIGLIVLILIVIGVPLIFTIKANSBACRDGLRAV 60  
 Db 1 mastsydcrypmedgdkrckllgiglivilivigvpliftikansseacrdgrav 60

Qy 61 MBCRNTHLQLQELTEAKQGKQDVEAQATCNHTWALMASLDAEKQGKVKVELEGET 120  
 Db 61 mectrvthlqlqelteaqkgfqdveaqatcnhtwalmaslaekaqgkveleget 120

Qy 121 TTLNKHQLQDASAEVERLRENQVLSVRIAKKYPPSSQDSSAAAPOLLVLLGSLALQ 180  
 Db 121 ttlnhkqlqdasaeeverlrenqvlsvriakkyppssqssaaapqlivlvgisalq 180

RESULT 11  
 AAB70697  
 ID AAB70697 standard; Protein; 180 AA.

XX  
 AC AAB70697;  
 XX  
 DT 18-MAY-2001 (first entry)  
 XX  
 DE Human HML.24 protein antigen SEQ ID NO:2.  
 XX  
 KW Human; HML.24 antigen expression potentiator; HML.24 protein antigen;  
 KW myeloma; Interferon alpha; Interferon gamma; IFN-alpha; IFN-gamma;  
 KW interferon regulatory factor 2; IRF-2; cytostatic; cytotoxic antibody;  
 KW multiple myeloma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200113940-A1.  
 XX  
 PR 01-MAR-2001.  
 XX  
 PR 22-AUG-2000; 2000WO-JP05617.  
 XX  
 PR 23-AUG-1999; 99JP-0236007.  
 XX  
 PR 16-FEB-2000; 2000JP-0038889.  
 XX  
 PA (CHUS ) CHUGAI SEIYAKU KK.  
 XX  
 PI Kosaka M, Ozaki S, Wakahara Y;  
 XX  
 DR WPI; 2001-202921/20.  
 XX  
 DR N-PSDB; AAT74792.  
 XX  
 PR HML.24 antigen expression potentiating agent containing interferon PR alpha or gamma or IRF-2 for treatment of myeloma -  
 XX  
 PS Claim 1; Page 55-56; 72pp; Japanese.  
 XX  
 CC The present invention describes an agent for potentiating the expression of HML.24 antigen in myeloma cells. The agent contains as an active component interferon (IFN) alpha or gamma, or interferon regulatory factor 2 (IRF-2), or a compound promoting the expression of IRF-2. Also described are: (1) drug compositions for the treatment of myeloma which contain the HML.24 antigen expression potentiating agent together with a cytotoxic antibody binding to HML.24 antigen; (2) screening compounds for their activity in potentiating the expression of IRF-2; and (3) kits for the treatment of myeloma using the HML.24 antigen expression potentiating agent and a cytotoxic antibody. The agent has cytosatic activity and can be used for the treatment of myeloma, especially of multiple myeloma. The present sequence represents the human HML.24 CC protein antigen.  
 XX  
 SQ Sequence 180 AA;

Query Match 100.0%; Score 889; DB 22; Length 180; Best Local Similarity 100.0%; Pred. 6.2e-78; Indels 0; Gaps 0; Matches 180; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 MASTSVDICRYPMEDGDKRCKLLGIGLIVLILIVIGVPLIFTIKANSBACRDGLRAV 60  
 Db 1 mastsydcrypmedgdkrckllgiglivilivigvpliftikansseacrdgrav 60

Qy 61 MBCRNTHLQLQELTEAKQGKQDVEAQATCNHTWALMASLDAEKQGKVKVELEGET 120  
 Db 61 mectrvthlqlqelteaqkgfqdveaqatcnhtwalmaslaekaqgkveleget 120

Qy 121 TTLNKHQLQDASAEVERLRENQVLSVRIAKKYPPSSQDSSAAAPOLLVLLGSLALQ 180  
 Db 121 ttlnhkqlqdasaeeverlrenqvlsvriakkyppssqssaaapqlivlvgisalq 180

RESULT 12  
 AAG73947  
 ID AAG73947 standard; Protein; 193 AA.

XX

- Eri-Jal 19 18:31:10 2002

AC AAG73947; AAW77292; ID AAW77292 standard; Protein; 180 AA.  
 XX DE 03-SEP-2001 (first entry)  
 XX DE Human colon cancer antigen protein SEQ ID NO:4711.  
 XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
 XX colorectal carcinoma; chromosome 19.  
 XX OS Homo sapiens.  
 XX PN WO200122920-A2.  
 XX PD 05-APR-2001.  
 XX PF 28-SEP-2000; 2000WO-US26524.  
 XX PR 29-SEP-1999; 99US-0157137.  
 XX PR 03-NOV-1999; 99US-0163280.  
 XX PA ( HUMA- ) HUMAN GENOME SCI INC.  
 XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;  
 XX WPI; 2001-235357/24.  
 XX N-PSDB; AAH33378.  
 PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides, useful for preventing, diagnosing and/or treating colorectal cancers -  
 PT Claim 11; Page 6512-6513; 9803pp; English.  
 XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon cancer-associated nucleic acid molecules (N) and proteins (P), where the proteins are collectively known as colon cancer antigens. The colon cancer antigens have cytostatic activity and can be used in gene therapy and vaccine production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression. For example, N and P may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of P by expressing inactive proteins or to supplement the patient's own production of P. Additionally, N may be used to produce the colon cancer-associated Ps, by inserting the nucleic acids into a host cell and culturing the cell to express the proteins. N and P can be used in the prevention, diagnosis and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204 and AAB7789 represent sequences used in the exemplification of the present invention.  
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were missing at time of publication, meaning no sequences are present for CC SEQ ID NO:1027 to 1052, 7921 and 7922.  
 XX Sequence 193 AA:  
 Query Match 99 2%; Score 889; DB 22; Length 193;  
 Best Local Similarity 100.0%; Pred. No. 6; 8e-78;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC QY 1 MASTSYDCKRPMEDGDKRCKLIGIGIYLILIVLGLPLIFTIKANSEACRDGLR 60  
 CC Db 1 mastsydckrpkmedgdckrligigilivlivilgpliftikansseacdrgir 60  
 CC SQ Sequence 180 AA;  
 Query Match 99 2%; Score 882; DB 19; Length 180;  
 Best Local Similarity 99.4%; Pred. No. 3e-77;  
 Matches 179; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CC QY 1 MASTSYDCKRPMEDGDKRCKLIGIGIYLILIVLGLPLIFTIKANSEACRDGLR 60  
 CC Db 1 mastsydckrpkmedgdckrligigilivlivilgpliftikansseacdrgir 60  
 CC SQ Sequence 180 AA;  
 Query Match 100.0%; Score 889; DB 22; Length 193;  
 Best Local Similarity 100.0%; Pred. No. 6; 8e-78;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC QY 1 MASTSYDCKRPMEDGDKRCKLIGIGIYLILIVLGLPLIFTIKANSEACRDGLR 60  
 CC Db 1 mastsydckrpkmedgdckrligigilivlivilgpliftikansseacdrgir 73  
 CC QY 61 MECRNWTHLQLQELTEAQKGFDQEADATCNYTMALMASDAEKQGQKKEELGEI 120  
 CC Db 61 mecnrvthlqlqelteaqkgfdqeadaqtcnhtwmalmasdaekqgqkkyvelegei 120  
 CC QY 121 TTLMHKLQDASAEVERLRENQVLVSRLAKDKKYPSSODSSAAAPQILIVLGLSALIQ 180  
 CC Db 121 ttlmhkldqdasaverlrenqvlvsrlakdkyypssodssaaapqilivlglslalq 180  
 CC QY 121 TTLMHKLQDASAEVERLRENQVLVSRLAKDKKYPSSODSSAAAPQILIVLGLSALIQ 180  
 CC Db 134 tlmhkldqdasaverlrenqvlvsrlakdkyypssodssaaapqilivlglslalq 193  
 RESULT 14  
 AAW36951  
 ID AAW36951 standard; Protein; 197 AA.  
 XX AC AAW36951;  
 XX DE 14-DEC-1998 (first entry)  
 XX DT 12-MAY-1998 (first entry)  
 XX DE Protein encoded by clone 0238\_1.  
 DE





Db	575	GAADLQNRIRLAESSLTRRLQESAAEVIALRDRDLERTAAS---RAQDVS	627	SQ	SEQUENCE	143 AA:	16386 MW;	613446D64B1C5402 CRC64;
RESULT	2			Query Match	12.0%	Score 107;	DB 16;	Length 143;
ID	Q9K6X4	PRELIMINARY;	PRT;	Best Local Similarity	22.1%	Pred. No. 0	13;	
ID	Q9K6X4			Matches	31;	Conservative	34;	Mismatches 49;
AC	Q9K6X4			Indels	26;	Gaps	4;	
DT	01-OCT-2000	(TREMBrel. 15, Created)		QY	23	LLGIGILVLILIVLGYPLIFTIKANSEACRDGLRAVMEC-----RNVTHILQQL 74		
DT	01-OCT-2000	(TREMBrel. 15, Last sequence update)		Db	6	LVNIMLILILIVLTVGVAVLIFVNLYFNNEDEODREPTIDEIQAQSYETEETNTLNSDF 65		
DT	01-DEC-2001	(TREMBrel. 19, Last annotation update)		QY	75	TEA-----OKGFQDVEAQATCNITVMAIMASLDAEKAQGQKVEELGEITTLNK 126		
DE	CELL WALL-BINDING PROTEIN			Db	66	VRARFLIHVNRNALQBVQKRDFQVNMIIRSLAGMDASQLSGADGIEKLRQ----- 118		
GN	B33600.			QY	127	LDASAEVERLRENQVLSV 146		
OS	Bacillus halodurans.			Db	119	LQD---DINALMHQEGSTVKKI 135		
OC	Bacteria; Firmicutes; Bacillus/clostridium group;							
OC	Bacillus/Staphylococcus group; Bacillus.							
RN	[1]	SEQUENCE FROM N.A.						
RX	STRAIN=C-125 / JCM 9153;							
RA	MEDLINE=20512582; PubMed=11058132;							
RA	Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,							
RA	Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,							
RA	Horikoshi K.;							
RT	"Complete genome sequence of the alkaliiphilic bacterium <i>Bacillus halodurans</i> and genomic sequence comparison with <i>Bacillus subtilis</i> ."							
RL	Nucleic Acids Res. 28:4317-4331(2000).							
DR	EMBL; AP001519; BAB07319.1; -.							
DR	MEPROS; M37 UPW; -.							
DR	InterPro; IPR02086; Peptidase_M37.							
DR	Pram; PR01557; Peptidase_M37; 1.							
DR	Complete proteome; Peptidase_M37; 1.							
SQ	SEQUENCE 461 AA; 50372 MW; 2918480CD67AFF3F CRC64;							
RESULT	3	Query Match	12.1%	Score 107.5;	DB 16;	Length 461;		
ID	Q9K43	Best Local Similarity	26.1%	Pred. No. 0	44;	Mismatches 31;		
ID	Q9K43	Matches	31;	AC	09Y6W2	Conservative	55;	Indels 5;
ID	Q9K43	AC		RT	09Y6W2;	PRT;	756 AA.	
DT	01-OCT-2000	(TREMBrel. 15, Created)		AC	Q9Y6W2;	PRELIMINARY;		
DT	01-OCT-2000	(TREMBrel. 15, Last sequence update)		DT	01-NOV-1999	(TREMBrel. 12, Created)		
DT	01-JUN-2001	(TREMBrel. 17, Last annotation update)		DT	01-NOV-1999	(TREMBrel. 12, Last sequence update)		
DE	FLAGELLAR PROTEIN REQUIRED FOR FLAGELLAR FORMATION.			DT	01-DEC-2001	(TREMBrel. 19, Last annotation update)		
GN	FIL1 OR BH247.			DE	HCR PROTEIN.			
OS	Bacillus halodurans.			DR	HCR.			
OC	Bacteria; Firmicutes; Bacillus/clostridium group;			OS	Homo sapiens (Human).			
OC	Bacillus/Staphylococcus group; Bacillus.			OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
RN	NCBI_TaxID=86655;			OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
RX	[1]	SEQUENCE FROM N.A.		OX	NCBI_TaxID=9606;			
RA	MEMLINE=2014706; PubMed=10545595;			RN	SEQUENCE FROM N.A.			
RA	Oka A., Tamai G., Tomizawa M., Ota M., Katsuyama Y., Makino S.,			RA	Shinoda T., Yoshimoto M., Mizuka M., Sasaki Y., Iwasita K.,			
RA	Kawakubo Y., Sugai J., Ozawa A., Okido M., Kimura M., Bahram S.,			RA	Inoko H.;			
RA	Inoko H.;			RT	"Association analysis using refined microsatellite markers localizes a susceptibility locus for psoriasis vulgaris within a 11kb segment telomeric to the HLA-C gene."			
RT	Hum. Mol. Genet. 8:2165-2170 (1999).			DR	Hum. Mol. Genet. 8:2165-2170 (1999).			
DR	EMBL; AB020331; BAA81890.1; -.			DR	EMBL; AB020331; BAA81890.1; -.			
SQ	SEQUENCE 756 AA; 86118 MW; 88FD5F858EF07601 CRC64;			DR	EMBL; AB020331; BAA81890.1; -.			
RESULT	4	Query Match	11.8%	Score 104.5;	DB 4;	Length 756;		
ID	Q9K43	Best Local Similarity	24.7%	Pred. No. 1	4;	Mismatches 39;		
ID	Q9K43	Matches	39;	AC	09Y6W2	Conservative	34;	Indels 48;
ID	Q9K43	AC		RT	09Y6W2;	PRT;	756 AA.	
DT	01-OCT-2000	(TREMBrel. 15, Created)		AC	Q9Y6W2;	PRELIMINARY;		
DT	01-OCT-2000	(TREMBrel. 15, Last sequence update)		DT	01-NOV-1999	(TREMBrel. 12, Created)		
DT	01-JUN-2001	(TREMBrel. 17, Last annotation update)		DT	01-NOV-1999	(TREMBrel. 12, Last sequence update)		
DE	FLAGELLAR PROTEIN REQUIRED FOR FLAGELLAR FORMATION.			DT	01-DEC-2001	(TREMBrel. 19, Last annotation update)		
GN	FIL1 OR BH247.			DE	HCR PROTEIN.			
OS	Bacillus halodurans.			DR	HCR.			
OC	Bacteria; Firmicutes; Bacillus/clostridium group;			OS	Homo sapiens (Human).			
OC	Bacillus/Staphylococcus group; Bacillus.			OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
RN	NCBI_TaxID=86655;			OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
RN	[1]	SEQUENCE FROM N.A.		OX	NCBI_TaxID=9606;			
RC	STRAIN=C-125 / JCM 9153;			RN	SEQUENCE FROM N.A.			
RX	MEMLINE=20512582; PubMed=11058132;			RC	STRAIN=C-125 / JCM 9153;			
RA	Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,			RA	Horikoshi K.;			
RA	Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,			RA	"Complete genome sequence of the alkaliiphilic bacterium <i>Bacillus halodurans</i> and genomic sequence comparison with <i>Bacillus subtilis</i> ."			
RA	RL	Nucleic Acids Res. 28:4317-4331(2000).		RA	EMBL; AP001515; BAB06166.1; -.			
RA	DE	Flagella; complete proteome.		RA	Flagella; complete proteome.			
RESULT	5	Query Match	12.0%	Score 107;	DB 16;	Length 143;		
ID	Q9NRK8	PRELIMINARY;	PRT;	Best Local Similarity	22.1%	Pred. No. 0	13;	
ID	Q9NRK8			Matches	31;	Conservative	34;	Mismatches 49;
ID	Q9NRK8			Indels	26;	Gaps	4;	
AC	Q9NRK8;			QY	23	LLGIGILVLILIVLGYPLIFTIKANSEACRDGLRAVMEC-----RNVTHILQQL 74		
DT	01-OCT-2000	(TREMBrel. 15, Created)		Db	6	LVNIMLILILIVLTVGVAVLIFVNLYFNNEDEODREPTIDEIQAQSYETEETNTLNSDF 65		
DT	01-OCT-2000	(TREMBrel. 15, Last sequence update)		QY	75	TEA-----OKGFQDVEAQATCNITVMAIMASLDAEKAQGQKVEELGEITTLNK 126		
DT	01-DEC-2001	(TREMBrel. 19, Last annotation update)		Db	66	VRARFLIHVNRNALQBVQKRDFQVNMIIRSLAGMDASQLSGADGIEKLRQ----- 118		
DE	CELL WALL-BINDING PROTEIN			QY	127	LDASAEVERLRENQVLSV 146		
GN	B33600.			Db	119	LQD---DINALMHQEGSTVKKI 135		
OS	Bacillus halodurans.							
OC	Bacteria; Firmicutes; Bacillus/clostridium group;							
OC	Bacillus/Staphylococcus group; Bacillus.							
RN	[1]	SEQUENCE FROM N.A.						
RX	STRAIN=C-125 / JCM 9153;							
RA	MEDLINE=20512582; PubMed=11058132;							
RA	Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,							
RA	Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,							
RA	Horikoshi K.;							
RT	"Complete genome sequence of the alkaliiphilic bacterium <i>Bacillus halodurans</i> and genomic sequence comparison with <i>Bacillus subtilis</i> ."							
RL	Nucleic Acids Res. 28:4317-4331(2000).							
DR	EMBL; AP001515; BAB06166.1; -.							
DR	Flagella; complete proteome.							







OC  
 OX  
 NC11\_TAXID=9913;  
 [1]  
 RP  
 SEQUENCE FROM N.A.  
 MEDLINE-97205278; PubMed=905278;  
 RA  
 Sawamura T., Kume N., Koyama T., Moriwaki H., Hoshikawa H., Aiba Y.,  
 RA  
 Tanaka T., Miwa S., Katsuma Y., Kita T., Masaki T.,  
 RT  
 "An endothelial receptor for oxidized low-density lipoprotein.",  
 RL  
 Nature 386:73-77(1997).  
 DR  
 EMBL: DB0494; BAA19005.1; -.  
 DR  
 HSSP: P20693; 1HLJ.  
 DR  
 Interpro: IPR001304; lectin\_c.  
 DR  
 Pfam: PF00059; lectin\_c; 1.  
 DR  
 SMART: SM0034; C1C8T; 1.  
 DR  
 PROSITE: PS50041; C\_TYPELECTIN\_2; 1.  
 KW  
 Receptor .  
 SQ  
 SEQUENCE 270 AA; 30892 MW; 6055B6891AD7053D CRC64;

Query Match 11.28; Score 99.5; DB 6; Length 270;  
 Best Local Similarity 25.9%; Pred. No. 1.1;  
 Matches 36; Conservative 25; Mismatches 41; Indels 37; Gaps 5;

Search completed: July 18, 2002, 18:44:32  
Job time: 135 sec

Job time: 135 sec

GenCore version 4.5 copyright (c) 1993 - 2000 Compugen Ltd.									
OM protein - protein search, using sw model									
Run on: July 18, 2002, 18:42:42 ; search time 11.94 seconds									
(without alignments)									
583.712 Million cell updates/sec									
Title: US-09-828-217-1	Sequence: 889	BUSSUM62	Gapop 10.0 , Gapext 0.5	105224	seqs, 38719550 residues	105224	hits satisfying chosen parameters:	105224	Post-processing: Minimum Match 0%
Scoring table: Gapop 10.0 , Gapext 0.5	Searched: 105224 seqs, 38719550 residues	Total number of hits satisfying chosen parameters:	105224	Minimum DB seq length: 0	Maximum DB seq length: 200000000	Maximum Match 0%	Post-processing: Minimum Match 100%	Post-processing: Maximum Match 100%	Post-processing: Listing first 45 summaries
Database : Swissprot_40:*	Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
SUMMARIES									
Result No.	Score	Query	Match length	DB	ID	Description	RESULT 1	STANDARD:	ALIGNMENTS
1	889	100.0	180	1	BST2_HUMAN	010589 homo sapien	BST2_HUMAN	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
2	102	11.5	550	1	KUCR_RAT	PI0176 rattus norv	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
3	101	11.4	422	1	KICR_MOUSE	PI05784 mus musculu	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
4	99.5	11.2	304	1	MNGL_MOUSE	PI49300 mus musculu	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
5	99	11.1	962	1	VDP_HUMAN	PI06763 homo sapien	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
6	98.5	11.1	1966	1	MSB_GAEL	PI02566 caenorhabdi	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
7	98	11.0	1938	1	MTH4_RABBIT	PI028641 oryctolagus	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
8	98	11.0	1938	1	MYS2_RABBIT	PI02567 caenorhabdi	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
9	97	10.9	286	1	PUG2_SCICO	PI22312 scirra copr	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
10	97	10.9	959	1	VDP_RAT	PI41542 ratulus norv	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
11	96.5	10.9	893	1	YNG2_CAEEL	PI03451 caenorhabdi	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
12	96	10.8	286	1	PUG1_SCICO	PI22311 scirra copr	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
13	95	10.7	1937	1	MTH8_HUMAN	PI13535 homo sapien	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
14	95	10.7	1972	1	MYHB_HUMAN	PI35749 homo sapien	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
15	95	10.7	1972	1	MYHB_RABBIT	PI35748 oryctolagus	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
16	94.5	10.6	1935	1	MYSS_CIPCA	PI09339 cyprinus ca	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
17	94.5	10.6	1935	1	PLEL_CIRGR	PI09155 criocerulus	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
18	94	10.6	1084	1	MYSS_RABBIT	PI02562 oryctolagus	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
19	93.5	10.5	1938	1	MYH6_RABBIT	PI02563 ratulus norv	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
20	93.5	10.5	1939	1	MYH6_MSMAU	PI13539 mesocricetus	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
21	93	10.5	848	1	MSP2_DIRIM	PI13392 dirofilaria	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
22	93	10.5	879	1	MSP2_OICVCO	PI001271 onchocerca	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
23	93	10.5	880	1	MSP2_BRUMA	PI001202 brucia mala	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
24	93	10.5	1972	1	MTHB_MOUSE	PI05659 acanthamoebe	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
25	92.5	10.4	1509	1	MYSN_ACACA	PI02566 mus musculu	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
26	92.5	10.4	1938	1	MYH6_MOUSE	PI10567 caenorhabdi	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
27	92	10.3	882	1	MSP2_CAEEL	PI11778 papio hamad	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
28	91.5	10.3	244	1	MTH7_PAPHA	PI042305 carassius a	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
29	91.5	10.3	428	1	FET2_CRAU	PI029616 gallus gall	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
30	91.5	10.3	1102	1	MYSC_CHICK	PI13540 mesocricetus	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
31	91.5	10.3	1934	1	MTH7_MSMAU	PI12883 homo sapien	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
32	91.5	10.3	1935	1	MYH7_HUMAN	PI02564 rattus norv	STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180
33	91.5	10.3	1935	1	MTH7_RAT		STANDARD:	PRT: 180 AA.	Query Match 100.0%; Score 889; DB 1; Length 180; Best Local Similarity 100.0%; Pred. No. 1.3e-63; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 MASTSYDVCVPMEDGDKRC.....SSAARQLLIVLIGLISALQ 180

Db	1	MASTSYDCRPMEDGDKRCKLKLIGIGILVLLIVLGVPLIFTIKANSACRDGLRAV 60	FT	CARBONYD	87	87	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	61	MCCRNVTHLLQOLETEAQKGFDQDVEAQATCNHTWALMASLDAEKAQGQKVELEGET 120	FT	CARBONYD	93	93	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	61	MCCRNVTHLLQOLETEAQKGFDQDVEAQATCNHTWALMASLDAEKAQGQKVELEGET 120	FT	CARBONYD	115	115	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	121	TTLNHKLQDASAVERLRENQVLSVRLADKAYVPSQDSSAAAPOLTVLIGSALLO 180	FT	CARBONYD	132	132	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	121	TTLNHKLQDASAVERLRENQVLSVRLADKAYVPSQDSSAAAPOLTVLIGSALLO 180	FT	CARBONYD	209	209	N-LINKED (GLCNAC. . .) (POTENTIAL).
US		Rattus norvegicus (Rat). Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. NCBI_TAXID=10116; [1]	FT	CARBONYD	255	255	N-LINKED (GLCNAC. . .) (POTENTIAL).
OC		SEQUENCE FROM N.A., AND SEQUENCE OF 83-104. MEDLINE=88227939; PubMed=2836387; [2]	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
RA		"Molecular cloning and sequencing of a cDNA for a carbohydrate binding receptor unique to rat Kupffer cells.;" J. Biol. Chem. 263:7487-7492(1988).	RT	KUCR RAT	STANDARD;	PRT;	550 AA.
RA		"Structure of the gene for a carbohydrate-binding receptor unique to rat Kupffer cells.;" J. Biol. Chem. 266:1850-1857(1991).	RT	KUCR RAT	STANDARD;	PRT;	550 AA.
RA		-I- FUNCTION: RECEPTOR WITH AN AFFINITY FOR GALACTOSE AND FUOSE. COULD BE INVOLVED IN ENDOCYTOSIS.	RT	KUCR RAT	STANDARD;	PRT;	550 AA.
CC		-I- SUBCELLULAR LOCATION: TYPE I membrane protein.	RT	KUCR RAT	STANDARD;	PRT;	550 AA.
CC		-I- TISSUE SPECIFICITY: KUPFER CELLS.	RT	KUCR RAT	STANDARD;	PRT;	550 AA.
CC		-I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.	RT	KUCR RAT	STANDARD;	PRT;	550 AA.
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement or send an email to license@ib-sib.ch).	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
CC		EMBL; M5532; AAA40892.1; -	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
DR		EMBL; J0374; AAA41472.1; -	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
DR		PIR; A28166; A28166.	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
DR		HSSP; P20693; 1HLJ.	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
DR		InterPro; IPR000117; Syntaxin.	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
DR		InterPro; IPR001304; Lectin.	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
DR		Pfam; PF00059; lectin_c; 1.	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
DR		SMART; SM00034; CLECT_1.	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
DR		PROSITE; PS00615; C_TYPE_LECTIN_1; 1.	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
DR		PROSITE; PS55041; C_TYPE_LECTIN_2; 1.	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
KW		Receptor; Transmembrane; Glycoprotein; lectin; Signal-anchor; Endocytosis.	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
FT		DOMAIN 1: CYTOSLAMIC (POTENTIAL). SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
FT		DOMAIN 70: 550 EXTRACELLULAR (POTENTIAL). C-TYPE LECTIN (SHORT FORM).	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
FT		DISULFID DISULFID BY SIMILARITY.	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
FT		DISULFID BY SIMILARITY.	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
CC		DISULFID BY SIMILARITY.	RP	KUCR RAT	STANDARD;	PRT;	550 AA.
Qy	46	IKANSEACRDLGRAVMECRNTHLLQOLETEAQKGFDQDVEAQATCNHTWALMASLDAEKAQGQKVELEGET 105	FT	CARBONYD	87	87	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	256	ISAEIQAMRDGMORAGE-----EMTSKKDLETLTAQIONANGHILQEQTDQIQL 305	FT	CARBONYD	115	115	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	106	KAQGOKKVELEGETITLNHKLQDASAVERLRENQVLSVRLADKAYVPSQDSSAAAPOLTVLIGSALLO 165	FT	CARBONYD	132	132	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	306	KAQ-LKSTSLSNSQEVNGKLDSSRELQTLRR-----LSDVSALKSNVQMLQSNLQAK 361	FT	CARBONYD	209	209	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	166	POLITVILGSA 177	FT	CARBONYD	255	255	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	362	AEVOSIKTGLEA 373	FT	CARBONYD	61104 MW;	61104 MW;	9358A6CF4C306270 CRC64;
Query Match 11.5%; Score 102; DB 1; Length 550; Best Local Similarity 28.0%; Pred. No. 0; 57; Mismatches 61; Indels 14; Gaps 3; Matches 37; Conservative 20; Mismatches 61; Indels 14; Gaps 3;							
Qy	46	IKANSEACRDLGRAVMECRNTHLLQOLETEAQKGFDQDVEAQATCNHTWALMASLDAEKAQGQKVELEGET 105	FT	CARBONYD	87	87	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	256	ISAEIQAMRDGMORAGE-----EMTSKKDLETLTAQIONANGHILQEQTDQIQL 305	FT	CARBONYD	115	115	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	106	KAQGOKKVELEGETITLNHKLQDASAVERLRENQVLSVRLADKAYVPSQDSSAAAPOLTVLIGSALLO 165	FT	CARBONYD	132	132	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	306	KAQ-LKSTSLSNSQEVNGKLDSSRELQTLRR-----LSDVSALKSNVQMLQSNLQAK 361	FT	CARBONYD	209	209	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	166	POLITVILGSA 177	FT	CARBONYD	255	255	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	362	AEVOSIKTGLEA 373	FT	CARBONYD	61104 MW;	61104 MW;	9358A6CF4C306270 CRC64;
Query Match 11.5%; Score 102; DB 1; Length 550; Best Local Similarity 28.0%; Pred. No. 0; 57; Mismatches 61; Indels 14; Gaps 3; Matches 37; Conservative 20; Mismatches 61; Indels 14; Gaps 3;							
Qy	46	IKANSEACRDLGRAVMECRNTHLLQOLETEAQKGFDQDVEAQATCNHTWALMASLDAEKAQGQKVELEGET 105	FT	CARBONYD	87	87	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	256	ISAEIQAMRDGMORAGE-----EMTSKKDLETLTAQIONANGHILQEQTDQIQL 305	FT	CARBONYD	115	115	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	106	KAQGOKKVELEGETITLNHKLQDASAVERLRENQVLSVRLADKAYVPSQDSSAAAPOLTVLIGSALLO 165	FT	CARBONYD	132	132	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	306	KAQ-LKSTSLSNSQEVNGKLDSSRELQTLRR-----LSDVSALKSNVQMLQSNLQAK 361	FT	CARBONYD	209	209	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	166	POLITVILGSA 177	FT	CARBONYD	255	255	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	362	AEVOSIKTGLEA 373	FT	CARBONYD	61104 MW;	61104 MW;	9358A6CF4C306270 CRC64;
Query Match 11.5%; Score 102; DB 1; Length 550; Best Local Similarity 28.0%; Pred. No. 0; 57; Mismatches 61; Indels 14; Gaps 3; Matches 37; Conservative 20; Mismatches 61; Indels 14; Gaps 3;							
Qy	46	IKANSEACRDLGRAVMECRNTHLLQOLETEAQKGFDQDVEAQATCNHTWALMASLDAEKAQGQKVELEGET 105	FT	CARBONYD	87	87	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	256	ISAEIQAMRDGMORAGE-----EMTSKKDLETLTAQIONANGHILQEQTDQIQL 305	FT	CARBONYD	115	115	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	106	KAQGOKKVELEGETITLNHKLQDASAVERLRENQVLSVRLADKAYVPSQDSSAAAPOLTVLIGSALLO 165	FT	CARBONYD	132	132	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	306	KAQ-LKSTSLSNSQEVNGKLDSSRELQTLRR-----LSDVSALKSNVQMLQSNLQAK 361	FT	CARBONYD	209	209	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	166	POLITVILGSA 177	FT	CARBONYD	255	255	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	362	AEVOSIKTGLEA 373	FT	CARBONYD	61104 MW;	61104 MW;	9358A6CF4C306270 CRC64;
Query Match 11.5%; Score 102; DB 1; Length 550; Best Local Similarity 28.0%; Pred. No. 0; 57; Mismatches 61; Indels 14; Gaps 3; Matches 37; Conservative 20; Mismatches 61; Indels 14; Gaps 3;							
Qy	46	IKANSEACRDLGRAVMECRNTHLLQOLETEAQKGFDQDVEAQATCNHTWALMASLDAEKAQGQKVELEGET 105	FT	CARBONYD	87	87	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	256	ISAEIQAMRDGMORAGE-----EMTSKKDLETLTAQIONANGHILQEQTDQIQL 305	FT	CARBONYD	115	115	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	106	KAQGOKKVELEGETITLNHKLQDASAVERLRENQVLSVRLADKAYVPSQDSSAAAPOLTVLIGSALLO 165	FT	CARBONYD	132	132	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	306	KAQ-LKSTSLSNSQEVNGKLDSSRELQTLRR-----LSDVSALKSNVQMLQSNLQAK 361	FT	CARBONYD	209	209	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	166	POLITVILGSA 177	FT	CARBONYD	255	255	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	362	AEVOSIKTGLEA 373	FT	CARBONYD	61104 MW;	61104 MW;	9358A6CF4C306270 CRC64;
Query Match 11.5%; Score 102; DB 1; Length 550; Best Local Similarity 28.0%; Pred. No. 0; 57; Mismatches 61; Indels 14; Gaps 3; Matches 37; Conservative 20; Mismatches 61; Indels 14; Gaps 3;							
Qy	46	IKANSEACRDLGRAVMECRNTHLLQOLETEAQKGFDQDVEAQATCNHTWALMASLDAEKAQGQKVELEGET 105	FT	CARBONYD	87	87	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	256	ISAEIQAMRDGMORAGE-----EMTSKKDLETLTAQIONANGHILQEQTDQIQL 305	FT	CARBONYD	115	115	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	106	KAQGOKKVELEGETITLNHKLQDASAVERLRENQVLSVRLADKAYVPSQDSSAAAPOLTVLIGSALLO 165	FT	CARBONYD	132	132	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	306	KAQ-LKSTSLSNSQEVNGKLDSSRELQTLRR-----LSDVSALKSNVQMLQSNLQAK 361	FT	CARBONYD	209	209	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	166	POLITVILGSA 177	FT	CARBONYD	255	255	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	362	AEVOSIKTGLEA 373	FT	CARBONYD	61104 MW;	61104 MW;	9358A6CF4C306270 CRC64;
Query Match 11.5%; Score 102; DB 1; Length 550; Best Local Similarity 28.0%; Pred. No. 0; 57; Mismatches 61; Indels 14; Gaps 3; Matches 37; Conservative 20; Mismatches 61; Indels 14; Gaps 3;							
Qy	46	IKANSEACRDLGRAVMECRNTHLLQOLETEAQKGFDQDVEAQATCNHTWALMASLDAEKAQGQKVELEGET 105	FT	CARBONYD	87	87	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	256	ISAEIQAMRDGMORAGE-----EMTSKKDLETLTAQIONANGHILQEQTDQIQL 305	FT	CARBONYD	115	115	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	106	KAQGOKKVELEGETITLNHKLQDASAVERLRENQVLSVRLADKAYVPSQDSSAAAPOLTVLIGSALLO 165	FT	CARBONYD	132	132	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	306	KAQ-LKSTSLSNSQEVNGKLDSSRELQTLRR-----LSDVSALKSNVQMLQSNLQAK 361	FT	CARBONYD	209	209	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	166	POLITVILGSA 177	FT	CARBONYD	255	255	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	362	AEVOSIKTGLEA 373	FT	CARBONYD	61104 MW;	61104 MW;	9358A6CF4C306270 CRC64;
Query Match 11.5%; Score 102; DB 1; Length 550; Best Local Similarity 28.0%; Pred. No. 0; 57; Mismatches 61; Indels 14; Gaps 3; Matches 37; Conservative 20; Mismatches 61; Indels 14; Gaps 3;							
Qy	46	IKANSEACRDLGRAVMECRNTHLLQOLETEAQKGFDQDVEAQATCNHTWALMASLDAEKAQGQKVELEGET 105	FT	CARBONYD	87	87	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	256	ISAEIQAMRDGMORAGE-----EMTSKKDLETLTAQIONANGHILQEQTDQIQL 305	FT	CARBONYD	115	115	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	106	KAQGOKKVELEGETITLNHKLQDASAVERLRENQVLSVRLADKAYVPSQDSSAAAPOLTVLIGSALLO 165	FT	CARBONYD	132	132	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	306	KAQ-LKSTSLSNSQEVNGKLDSSRELQTLRR-----LSDVSALKSNVQMLQSNLQAK 361	FT	CARBONYD	209	209	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	166	POLITVILGSA 177	FT	CARBONYD	255	255	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	362	AEVOSIKTGLEA 373	FT	CARBONYD	61104 MW;	61104 MW;	9358A6CF4C306270 CRC64;
Query Match 11.5%; Score 102; DB 1; Length 550; Best Local Similarity 28.0%; Pred. No. 0; 57; Mismatches 61; Indels 14; Gaps 3; Matches 37; Conservative 20; Mismatches 61; Indels 14; Gaps 3;							
Qy	46	IKANSEACRDLGRAVMECRNTHLLQOLETEAQKGFDQDVEAQATCNHTWALMASLDAEKAQGQKVELEGET 105	FT	CARBONYD	87	87	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	256	ISAEIQAMRDGMORAGE-----EMTSKKDLETLTAQIONANGHILQEQTDQIQL 305	FT	CARBONYD	115	115	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	106	KAQGOKKVELEGETITLNHKLQDASAVERLRENQVLSVRLADKAYVPSQDSSAAAPOLTVLIGSALLO 165	FT	CARBONYD	132	132	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	306	KAQ-LKSTSLSNSQEVNGKLDSSRELQTLRR-----LSDVSALKSNVQMLQSNLQAK 361	FT	CARBONYD	209	209	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	166	POLITVILGSA 177	FT	CARBONYD	255	255	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	362	AEVOSIKTGLEA 373	FT	CARBONYD	61104 MW;	61104 MW;	9358A6CF4C306270 CRC64;
Query Match 11.5%; Score 102; DB 1; Length 550; Best Local Similarity 28.0%; Pred. No. 0; 57; Mismatches 61; Indels 14; Gaps 3; Matches 37; Conservative 20; Mismatches 61; Indels 14; Gaps 3;							
Qy	46	IKANSEACRDLGRAVMECRNTHLLQOLETEAQKGFDQDVEAQATCNHTWALMASLDAEKAQGQKVELEGET 105	FT	CARBONYD	87	87	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	256	ISAEIQAMRDGMORAGE-----EMTSKKDLETLTAQIONANGHILQEQTDQIQL 305	FT	CARBONYD	115	115	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	106	KAQGOKKVELEGETITLNHKLQDASAVERLRENQVLSVRLADKAYVPSQDSSAAAPOLTVLIGSALLO 165	FT	CARBONYD	132	132	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	306	KAQ-LKSTSLSNSQEVNGKLDSSRELQTLRR-----LSDVSALKSNVQMLQSNLQAK 361	FT	CARBONYD	209	209	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	166	POLITVILGSA 177	FT	CARBONYD	255	255	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	362	AEVOSIKTGLEA 373	FT	CARBONYD	61104 MW;	61104 MW;	9358A6CF4C306270 CRC64;
Query Match 11.5%; Score 102; DB 1; Length 550; Best Local Similarity 28.0%; Pred. No. 0; 57; Mismatches 61; Indels 14; Gaps 3; Matches 37; Conservative 20; Mismatches 61; Indels 14; Gaps 3;							
Qy	46	IKANSEACRDLGRAVMECRNTHLLQOLETEAQKGFDQDVEAQATCNHTWALMASLDAEKAQGQKVELEGET 105	FT	CARBONYD	87	87	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	256	ISAEIQAMRDGMORAGE-----EMTSKKDLETLTAQIONANGHILQEQTDQIQL 305	FT	CARBONYD	115	115	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	106	KAQGOKKVELEGETITLNHKLQDASAVERLRENQVLSVRLADKAYVPSQDSSAAAPOLTVLIGSALLO 165	FT	CARBONYD	132	132	N-LINKED (GLCNAC. . .) (POTENTIAL).
Db	306	KAQ-LKSTSLSNSQEVNGKLDSSRELQTLRR-----LSDVSALKSNVQMLQSNLQAK 361	FT	CARBONYD	209	209	N-LINKED (GLCNAC. . .) (POTENTIAL).
Qy	166	POLITVILGSA 177	FT	CARBONYD			





use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; J01050; AAA28124.1; -.

DR PIR; A02992; MRKW.

DR PIR; P08799; IINND.

DR InterPro; IPR004009; Myosin\_N.

DR InterPro; IPR002928; Myosin\_tail.

DR Pfam; PF01576; Myosin\_head; 1.

DR PRINTS; PRO0193; MYOSINHEAVY.

DR PRODOM; P0000355; myosin\_head; 1.

DR SMART; SM00242; MYSC; 1.

KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding; ATP-binding; Methylation; Alkylation; Multigene family.

FT DOMAIN 851 1966 COILED COIL (POTENTIAL).

FT DOMAIN 851 1966 ALPHA-HELICAL TAILPIECE (S2).

FT DOMAIN 1165 1966 HINGE.

FT DOMAIN 1165 1966 LIGHT MERO-MYOSIN (LMM).

FT NP\_BIND 177 184 ATP (BY SIMILARITY).

FT DOMAIN 665 687 ACTIN-BINDING.

FT DOMAIN 769 783 ACTIN-BINDING.

FT MOD\_RES 128 128 METHYLATION (TRI-) (POTENTIAL).

FT MOD\_RES 705 705 ALKYLATION (SH-2).

FT 715 715 E -> R (IN REF. 2).

FT 1337 1337 I -> L (IN REF. 2).

FT CONFFLICT 1880 1880 SEQUENCE 1966 AA; 225125 MW; B66F0BB2FE27B67F CRC64;

Query Match 11.1%; Score 98.5; DB 1; Length 1966; Best Local Similarity 24.7%; Pred. No. 4; Gaps 5; Matches 36; Conservative 21; Mismatches 58; Indels 31; Gaps 5;

QY 47 KANSEACRGIRAVECRNTHLQLQELTEAQKGQDVEAQATCNHTVMAA 106

Db 1811 EAAALKGGKKVIAKLEQVRVLESELDEQQRRQDANKNLGRADRRVRELQVDEK 1870

QY 107 AQQ-----GOKVQELEGETTLN-----HKLQDASAEVERLREN 141

Db 1871 KNNFRQLQDLIDKQOKLKTOKKQVEEAE-ELANLNQKQYKQTHOLEDAERRAD-QAEN 1927

QY 142 QVLSTRI---ADKKVYPSSQDSSA 164

QY 1928 SLSKMRSKSRASASVAPGLOSSASA 1953

RESULT 7

MYH4\_RABIT

ID MH4\_RABIT STANDARD; PRT; 1938 AA.

AC 028641;

DT 16-OCT-2001 (Rel. 40 Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Myosin heavy chain, skeletal muscle, juvenile.

OS Oryctolagus cuniculus (Rabbit).

OC Bikaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

NCBI-Taxid-9986;

RN [1] SEQUENCE FROM N A.

RC STRAIN-NEW ZEALAND WHITE; TISSUE=Skeletal muscle;

RA Maeda K., Hostinova E., Roesch-Kleinckauf A., Schuster H., Gasperik J.,

RA Wittinghofer A.;

RT "Isolation, sequencing of myosin heavy chain cDNA from rabbit skeletal muscle and a novel cosynthesis of S-1 fragment with the

RT essential and regulatory light chains.";

RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: MUSCLE CONTRACTION.

CC -!- SUBUNITS: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (HHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).

CC -!- SUBCELLULAR LOCATION: Thick filaments.

CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.

CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATTER BE SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHARED SUBFRAGMENT (S2).

CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.

CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.

CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATTER BE SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHARED SUBFRAGMENT (S2).

CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.

CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U32574; AAA74199.1; -.

DR HSSP; P08799; IINND.

DR InterPro; IPR000048; IQ.

DR InterPro; IPR04009; Myosin\_N.

DR InterPro; IPR002928; Myosin\_tail.

DR InterPro; IPR001609; myosin\_head.

DR Pfam; PF00612; IQ; 2.

DR Pfam; PF00063; myosin\_head; 1.

DR Pfam; PF01576; Myosin\_N; 1.

DR PRINTS; PRO0193; MYOSINHEAVY.

DR PRODOM; P0000355; myosin\_head; 1.

DR SMART; SM00015; IQ; 1.

DR SMART; SM00242; MYSC; 1.

DR PROSITE; PS50096; IQ; 1.

KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding; Calmodulin-binding; ATP-binding; Methylation; Alkylation; Multigene family.

FT DOMAIN 784 813 IQ.

FT DOMAIN 842 1938 COILED COIL (POTENTIAL).

FT NP\_BIND 179 186 ATP (POTENTIAL).

FT DOMAIN 658 680 ACTIN BINDING (BY SIMILARITY).

FT DOMAIN 760 774 ACTIN-BINDING (BY SIMILARITY).

FT MOD\_RES 35 35 METHYLATION (MONO-) (BY SIMILARITY).

FT MOD\_RES 130 130 METHYLATION (TRI-) (BY SIMILARITY).

FT MOD\_RES 552 552 METHYLATION (TRI-) (BY SIMILARITY).

FT MOD\_RES 756 756 METHYLATION (TRI-) (BY SIMILARITY).

FT MOD\_RES 698 698 ALKYLATION (SH-1) (BY SIMILARITY).

FT MOD\_RES 708 708 ALKYLATION (SH-2) (BY SIMILARITY).

SQ SEQUENCE 1938 AA; 223064 MW; D8A8A2EC5B182626 CRC64;

Query Match 11.0%; Score 98; DB 1; Length 1938; Best Local Similarity 26.1%; Pred. No. 4; Gaps 45; Matches 37; Conservative 20; Mismatches 45; Indels 40; Gaps 5; Gaps 5;

QY 47 KANSEACR-----DGLRAVECRNTHLQLQELTEAQKGQDVEAQATCNHTVMAA 99

Db 1366 KANSEFAQWQKTYETDAQRTTELEAKKLAQRLDAE-EHVBANAKC----- 1414

QY 100 ASLDAAKAQGQKVELEGETTLN-----HKLQDASAEVERLRL 138

Db 1415 ASLEKTRQRLQNEVEDLMDIDVERTNAACALDKKQRNFDKILAEWKHKYETHAELEASQ 1474

QY 139 RENQVLSTRIAD-KKYPSSQD 159

Db 1475 KESRSISTEVVKVKNAYEESID 1496

RESULT 8  
 MISS\_CABEL DR EMBL: 271266; CAA95806.1; JOINED.  
 ID MYSD\_CABEL STANDARD; PRT: 1938 AA.  
 DR PIR: S03772; MWKWI.  
 DR HSSP; P08799; IMND.  
 DR WormPep; R05C7.10; CED6253.  
 DR InterPro; IPR004009; Myosin\_N.  
 DR InterPro; IPR00228; Myosin\_tail.  
 DR Pfam; PF00063; myosin\_head; 1.  
 DR Pfam; PF02736; Myosin\_N; 1.  
 DR Pfam; PF01576; Myosin\_tail; 1.  
 DR PRINS; PRO0193; MYOSINHEAVY.  
 DR PRODOM; PD000355; myosin\_head; 1.  
 DR SMART; SM00342; MYSC; 1.  
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KW ATP-binding; Methylation; Alkylation; Multigene family.  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE=89178677; PubMed=2926820;  
 RT "sequence analysis of the complete *Caenorhabditis elegans* myosin  
 heavy chain gene family";  
 RT heavy chain gene family; J. Mol. Biol. 205:603-613(1989).  
 RL [2]  
 RN SEQUENCE OF 34-1795 FROM N.A.  
 RX MEDLINE=83273600; PubMed=6576334;  
 RA Karn J., Maruyama T.N., Krause M., Karn J.;  
 RT "Protein structural domains in the *Caenorhabditis elegans* unc-54  
 myosin heavy chain gene are not separated by introns";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).  
 RN [3]  
 SEQUENCE OF 115-365 AND 1422-1763 FROM N.A.  
 RX MEDLINE=85201409; PubMed=3488374;  
 RA Karn J., Dibb N.J., Miller D.M.;  
 RT "Cloning nematode myosin genes";  
 RL Cell Muscle Motil. 6:185-237(1985).  
 RN [4]  
 SEQUENCE FROM N.A.  
 RP STRAIN=BRISTOL N2;  
 RA Gardner A., McMurry A.;  
 RL Submitted (AAR-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: MUSCLE CONTRACTION.  
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 HEAVY CHAIN SUBUNITS (HMC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.  
 CC -1- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN THE PHARYNGEAL MUSCLE.  
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1- TWO CYSTINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 MEROMYOSIN (LHM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN  
 CC C. ELEGANS.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC

RESULT 9  
 P092\_SCICO PRT: 286 AA.  
 ID P192\_SCICO STANDARD; PRT: 286 AA.  
 AC P22312;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE PUFF\_II/9-2 protein precursor.  
 GN 11/9-2.  
 OS Sciaridae; Brachycera.  
 OC Sciaridae; Brachycera.  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Sciaroidea;  
 OC Sciaridae; Brachycera.  
 DR NCBI\_TaxID=38358;  
 RN [1]  
 SEQUENCE FROM N.A.  
 EMBL: X08065; CAA30854.1; -.  
 DR EMBL; M3732; AAC28119.1; -.  
 DR EMBL; M37234; AAC38120.1; -.  
 DR EMBL; 271266; CAA95848.1; -.  
 DR EMBL; 271261; CAA95848.1; JOINED.  
 DR EMBL; 271261; CAA95806.1; -.



DE HYPOTHETICAL 100.0 kDa protein M01A8.2 in chromosome III.  
 GN M01A8.2  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderaidae; Caenorhabditidae.  
 OC NCBI\_TaxID=6239;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RP STRAIN=BRISTOL N2;  
 RX MEDLINE=94150718; PubMed=7966398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Berks M., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laisser N.,  
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifkin L., Roopa D., Showman R.,  
 RA Sims M., Smalton N., Smith A., Smith M., Sonnhammer E., Staden R.,  
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
 RA Wohldman P., Watson A., Weinstock L., Wilkinson-Sproat J.,  
 RA RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RL elegans", Nature 368:32-38(1994).  
 CC -!- SIMILARITY: CONTAINS 1 CAP-GLY DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC use by non profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC ---  
 DR EMBL; Z27081; CAA81607.1; -.  
 DR PIR; S40998; S40998.  
 DR WormPep; M01A8.2; CE03491.  
 DR InterPro; IPR000938; CAP-GLY.  
 DR Pfam; PF01302; CAP-GLY; 1.  
 DR PROSITE; PS00845; CAP-GLY\_1; 1.  
 DR PROSITE; PS50245; CAP-GLY\_2; 1.  
 KW HYPOTHETICAL PROTEIN; COILED COIL.  
 FT DOMAIN 522 696 COILED COIL (POTENTIAL).  
 FT DOMAIN 729 756 COILED COIL (POTENTIAL).  
 SEQUENCE 893 AA: 99997 MN: 464F2962B36C28B1 CRC64;

Query Match 10.9%; Score 96.5; DB 1; Length 893;  
 Best Local Similarity 24.8%; Pred. No. 2.6;  
 Matches 32; Conservative 29; Mismatches 39; Indels 29; Gaps 4;

QY 48 ANEACACRGLRATMECRNTHLILQ-QLITEAQKGFQDV---EAQATCNHTVIMALMAS 101  
 DB 566 SNOOVRINHANAY-ESLOKTHETQIAEKNEKEFERNFEERARREAEVCMNNHRQVVA 624  
 QY 102 LDDEKAQGOKKVBLE-----GEITTLNKHQDASAVERIRR 139  
 DB 625 LDEKISEAEKQCOLQNLNVKVKVQALANDCDHQNQMLTKEISLQTALEMSEKELRQ 684  
 QY 140 ENQYLSVRI 148  
 DB 685 KNQNLSLQV 693

RESULT 12  
 PUBL\_SCICO STANDARD; PRT; 286 AA.  
 AC P23311;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 RL t;

DE PUFF II/9-1 protein precursor.  
 GN IT/9-1.  
 OS Sciaracoprophila (Fungus gnat).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Sciaroidea;  
 OC Sciaridae; Brachysia.  
 OC NCBI\_TaxID=38358;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RP STRAIN=6980;  
 RX MEDLINE=90133907; PubMed=2614832;  
 RA Dibartolomei S.M., Gerbi S.A.;  
 RT "Molecular characterization of DNA puff II/9A genes in Sciaracoprophila.", J. Mol. Biol. 210:531-540(1989).  
 CC -!- MISCELLANEOUS: THE DNA PUFF II/9 PROTEINS HAVE A PROPOSED  
 CC -!- INTERMOLECULAR COILED COIL STABILITY WITH POSSIBLY INTERMOLECULAR  
 CC DISULFIDE BRIDGES FORMED BY NUMEROUS CYSTEINE RESIDUES IN POSITION  
 CC D OF THE HEPTAD REPEAT.  
 CC -!- SIMILARITY: 76% IDENTICAL TO THE PUFF II/9-2 PROTEIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC ---  
 DR EMBL; X51680; CAA35983.1; -.  
 DR PIR; S07532; S07532.  
 KW Signal; Coiled coil; Glycoprotein.  
 FT SIGNAL 1 19 OR 21 (POTENTIAL).  
 FT CHAIN 20 286 PUFF II/9-1 PROTEIN.  
 FT DOMAIN 61 235 HELICAL (POTENTIAL).  
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SEQUENCE 286 AA: 32034 MN: AA6A7B5F191BB1D CRC64;

Query Match 10.8%; Score 96; DB 1; Length 286;  
 Best Local Similarity 30.0%; Pred. No. 0.84;  
 Matches 24; Conservative 16; Mismatches 36; Indels 4; Gaps 1;

QY 59 AVMECRNTHLILQELTEAQKGFQDVRAAQATCNHTVIMALMASLDAEKAQSKKVERLEG 118  
 DB 102 ALCECOKNSELLIKOTIQELKLAQTKELANCKEA---LANCKAENAKLKKIEELNC 157  
 QY 119 EITTLNKHQDASAVERIRR 138  
 DB 158 TIQLOQEELEQRARERDQ 177

RESULT 13  
 MYB\_HUMAN ID M08\_HUMAN STANDARD; PRT; 1937 AA.  
 AC P13535; Q14910;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 01-NOV-2001 (Rel. 40, Last annotation update)  
 DE Myosin heavy chain, skeletal muscle, perinatal (MyHC-perinatal).  
 MYB\_HUMAN  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Cetartiodactyla; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RP TISSUE-Skeletal muscle;  
 RX MEDLINE=9033631; PubMed=2373371;  
 RA Karisch-Mizrahi I., Reghali R., Shows T.B. Jr., Leinwand L.A.;  
 RT "Generation of a full-length human perinatal myosin heavy-chain-encoding cDNA.", Gene 89:289-294(1990).

RN [2] SEQUENCE FROM N.A.  
 RE DR SMART; SM00015; IQ; 1.  
 RC DR SMART; SM00242; MYSC; 1.  
 RX PROSITE; PS50096; IQ; 1;  
 RA Myosin: Muscle protein; coiled coil; Thick filament; Actin-binding;  
 RA Julian E.H., Kelly A.M., Pompidou A.J., Hoffman R., Schiaffino S.,  
 RA Stedman H., Rubinstein N.A.;  
 RT "characterization of a human perinatal myosin heavy-chain  
 transcript.";  
 RT Bur. J. Biochem. 230:1001-1006(1995).  
 RN [3] SEQUENCE OF 502-1937 FROM N.A.  
 RE TISSUE-Skeletal muscle;  
 RC MEDLINE=90235862; PubMed=6691980;  
 RX Bober E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,  
 RA Arnold H.H.;  
 RT "Identification of three developmentally controlled isoforms of human  
 myosin heavy chains.";  
 RT Eur. J. Biochem. 189:55-65(1990).  
 RN [4] SEQUENCE OF 860-1937 FROM N.A.  
 RE MEDLINE=89334168; PubMed=2715179;  
 RA Feghali R., Leinwand L.A.;  
 RA "Molecular genetic characterization of a developmentally regulated  
 human perinatal myosin heavy chain.";  
 J. Cell Biol. 108:1791-1797(1989).  
 RN [5] SEQUENCE OF 1-46 FROM N.A.  
 RP Esser K., Tidhar A., Myszkowski M.;  
 RA "Isolation and characterization of the human perinatal MHC promoter.";  
 RT RL Submitted (MAY-1998) to the EMBL/GeneBank/DBJ databases.  
 CC CC  
 CC -!- FUNCTION: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC  
 CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -!- DOMAIN: THE RODLIKE TAIL SUBSEQUENCE IS HIGHLY REPEATITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MERO-MYOSIN (LMM) AND 1 HEAVY MERO-MYOSIN (HMM). IT CAN LAYER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL; M36759; AAC17195.1; -.  
 DR EMBL; 238133; CAB86293.1; -.  
 DR EMBL; X51592; CAB35941.1; -.  
 DR EMBL; M3250; AAC36346.1; -.  
 DR EMBL; AF067143; AAC21557.1; -.  
 DR PIR; A30220; A30220.  
 DR HSSP; P08759; ILVK.  
 DR MIM; 160741; -.  
 DR InterPro; IPR000448; IQ.  
 DR InterPro; IPR000099; Myosin\_N.  
 DR InterPro; IPR00228; Myosin\_tail.  
 DR InterPro; IPR001609; myosin\_head.  
 DR Pfam; PF00612; IQ; 1.  
 DR Pfam; PF00063; myosin\_head; 1.  
 DR Pfam; PF02736; Myosin\_N; 1.  
 DR Pfam; PF01576; Myosin\_tail; 1.  
 DR PRINTS; PR00193; MYOSINHEAVY.  
 DR PRODOM; PD000355; myosin\_head; 1.  
 RX MEDLINE=99156230; PubMed=10048485;

DR SMART; SM00015; IQ; 1.  
 DR SMART; SM00242; MYSC; 1.  
 DR PROSITE; PS50096; IQ; 1;  
 RA Myosin: Muscle protein; coiled coil; Thick filament; Actin-binding;  
 RA Julian E.H., Kelly A.M., Pompidou A.J., Hoffman R., Schiaffino S.,  
 RA Stedman H., Rubinstein N.A.;  
 RA "characterization of a human perinatal myosin heavy-chain  
 transcript.";  
 RA Bur. J. Biochem. 230:1001-1006(1995).  
 RA Feghali R., Leinwand L.A.;  
 RA "Molecular genetic characterization of a developmentally regulated  
 human perinatal myosin heavy chain.";  
 RA Eur. J. Biochem. 189:55-65(1990).  
 RA "Isolation and characterization of the human perinatal MHC promoter.";  
 RA Submitted (MAY-1998) to the EMBL/GeneBank/DBJ databases.  
 CC CC  
 CC -!- FUNCTION: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC  
 CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -!- DOMAIN: THE RODLIKE TAIL SUBSEQUENCE IS HIGHLY REPEATITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MERO-MYOSIN (LMM) AND 1 HEAVY MERO-MYOSIN (HMM). IT CAN LAYER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL; M36759; AAC17195.1; -.  
 DR EMBL; 238133; CAB86293.1; -.  
 DR EMBL; X51592; CAB35941.1; -.  
 DR EMBL; M3250; AAC36346.1; -.  
 DR EMBL; AF067143; AAC21557.1; -.  
 DR PIR; A30220; A30220.  
 DR HSSP; P08759; ILVK.  
 DR MIM; 160741; -.  
 DR InterPro; IPR000448; IQ.  
 DR InterPro; IPR000099; Myosin\_N.  
 DR InterPro; IPR00228; Myosin\_tail.  
 DR InterPro; IPR001609; myosin\_head.  
 DR Pfam; PF00612; IQ; 1.  
 DR Pfam; PF00063; myosin\_head; 1.  
 DR Pfam; PF02736; Myosin\_N; 1.  
 DR Pfam; PF01576; Myosin\_tail; 1.  
 DR PRINTS; PR00193; MYOSINHEAVY.  
 DR PRODOM; PD000355; myosin\_head; 1.  
 RX MEDLINE=99156230; PubMed=10048485;

DR SMART; SM00015; IQ; 1.  
 DR SMART; SM00242; MYSC; 1.  
 DR PROSITE; PS50096; IQ; 1;  
 RA Myosin: Muscle protein; coiled coil; Thick filament; Actin-binding;  
 RA Julian E.H., Kelly A.M., Pompidou A.J., Hoffman R., Schiaffino S.,  
 RA Stedman H., Rubinstein N.A.;  
 RA "characterization of a human perinatal myosin heavy-chain  
 transcript.";  
 RA Bur. J. Biochem. 230:1001-1006(1995).  
 RA Feghali R., Leinwand L.A.;  
 RA "Molecular genetic characterization of a developmentally regulated  
 human perinatal myosin heavy chain.";  
 RA Eur. J. Biochem. 189:55-65(1990).  
 RA "Isolation and characterization of the human perinatal MHC promoter.";  
 RA Submitted (MAY-1998) to the EMBL/GeneBank/DBJ databases.  
 CC CC  
 CC -!- FUNCTION: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC  
 CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -!- DOMAIN: THE RODLIKE TAIL SUBSEQUENCE IS HIGHLY REPEATITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MERO-MYOSIN (LMM) AND 1 HEAVY MERO-MYOSIN (HMM). IT CAN LAYER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL; M36759; AAC17195.1; -.  
 DR EMBL; 238133; CAB86293.1; -.  
 DR EMBL; X51592; CAB35941.1; -.  
 DR EMBL; M3250; AAC36346.1; -.  
 DR EMBL; AF067143; AAC21557.1; -.  
 DR PIR; A30220; A30220.  
 DR HSSP; P08759; ILVK.  
 DR MIM; 160741; -.  
 DR InterPro; IPR000448; IQ.  
 DR InterPro; IPR000099; Myosin\_N.  
 DR InterPro; IPR00228; Myosin\_tail.  
 DR InterPro; IPR001609; myosin\_head.  
 DR Pfam; PF00612; IQ; 1.  
 DR Pfam; PF00063; myosin\_head; 1.  
 DR Pfam; PF02736; Myosin\_N; 1.  
 DR Pfam; PF01576; Myosin\_tail; 1.  
 DR PRINTS; PR00193; MYOSINHEAVY.  
 DR PRODOM; PD000355; myosin\_head; 1.  
 RX MEDLINE=99156230; PubMed=10048485;

DR SMART; SM00015; IQ; 1.  
 DR SMART; SM00242; MYSC; 1.  
 DR PROSITE; PS50096; IQ; 1;  
 RA Myosin: Muscle protein; coiled coil; Thick filament; Actin-binding;  
 RA Julian E.H., Kelly A.M., Pompidou A.J., Hoffman R., Schiaffino S.,  
 RA Stedman H., Rubinstein N.A.;  
 RA "characterization of a human perinatal myosin heavy-chain  
 transcript.";  
 RA Bur. J. Biochem. 230:1001-1006(1995).  
 RA Feghali R., Leinwand L.A.;  
 RA "Molecular genetic characterization of a developmentally regulated  
 human perinatal myosin heavy chain.";  
 RA Eur. J. Biochem. 189:55-65(1990).  
 RA "Isolation and characterization of the human perinatal MHC promoter.";  
 RA Submitted (MAY-1998) to the EMBL/GeneBank/DBJ databases.  
 CC CC  
 CC -!- FUNCTION: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC  
 CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -!- DOMAIN: THE RODLIKE TAIL SUBSEQUENCE IS HIGHLY REPEATITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MERO-MYOSIN (LMM) AND 1 HEAVY MERO-MYOSIN (HMM). IT CAN LAYER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL; M36759; AAC17195.1; -.  
 DR EMBL; 238133; CAB86293.1; -.  
 DR EMBL; X51592; CAB35941.1; -.  
 DR EMBL; M3250; AAC36346.1; -.  
 DR EMBL; AF067143; AAC21557.1; -.  
 DR PIR; A30220; A30220.  
 DR HSSP; P08759; ILVK.  
 DR MIM; 160741; -.  
 DR InterPro; IPR000448; IQ.  
 DR InterPro; IPR000099; Myosin\_N.  
 DR InterPro; IPR00228; Myosin\_tail.  
 DR InterPro; IPR001609; myosin\_head.  
 DR Pfam; PF00612; IQ; 1.  
 DR Pfam; PF00063; myosin\_head; 1.  
 DR Pfam; PF02736; Myosin\_N; 1.  
 DR Pfam; PF01576; Myosin\_tail; 1.  
 DR PRINTS; PR00193; MYOSINHEAVY.  
 DR PRODOM; PD000355; myosin\_head; 1.  
 RX MEDLINE=99156230; PubMed=10048485;

DR SMART; SM00015; IQ; 1.  
 DR SMART; SM00242; MYSC; 1.  
 DR PROSITE; PS50096; IQ; 1;  
 RA Myosin: Muscle protein; coiled coil; Thick filament; Actin-binding;  
 RA Julian E.H., Kelly A.M., Pompidou A.J., Hoffman R., Schiaffino S.,  
 RA Stedman H., Rubinstein N.A.;  
 RA "characterization of a human perinatal myosin heavy-chain  
 transcript.";  
 RA Bur. J. Biochem. 230:1001-1006(1995).  
 RA Feghali R., Leinwand L.A.;  
 RA "Molecular genetic characterization of a developmentally regulated  
 human perinatal myosin heavy chain.";  
 RA Eur. J. Biochem. 189:55-65(1990).  
 RA "Isolation and characterization of the human perinatal MHC promoter.";  
 RA Submitted (MAY-1998) to the EMBL/GeneBank/DBJ databases.  
 CC CC  
 CC -!- FUNCTION: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC  
 CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -!- DOMAIN: THE RODLIKE TAIL SUBSEQUENCE IS HIGHLY REPEATITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MERO-MYOSIN (LMM) AND 1 HEAVY MERO-MYOSIN (HMM). IT CAN LAYER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL; M36759; AAC17195.1; -.  
 DR EMBL; 238133; CAB86293.1; -.  
 DR EMBL; X51592; CAB35941.1; -.  
 DR EMBL; M3250; AAC36346.1; -.  
 DR EMBL; AF067143; AAC21557.1; -.  
 DR PIR; A30220; A30220.  
 DR HSSP; P08759; ILVK.  
 DR MIM; 160741; -.  
 DR InterPro; IPR000448; IQ.  
 DR InterPro; IPR000099; Myosin\_N.  
 DR InterPro; IPR00228; Myosin\_tail.  
 DR InterPro; IPR001609; myosin\_head.  
 DR Pfam; PF00612; IQ; 1.  
 DR Pfam; PF00063; myosin\_head; 1.  
 DR Pfam; PF02736; Myosin\_N; 1.  
 DR Pfam; PF01576; Myosin\_tail; 1.  
 DR PRINTS; PR00193; MYOSINHEAVY.  
 DR PRODOM; PD000355; myosin\_head; 1.  
 RX MEDLINE=99156230; PubMed=10048485;

RA Nagase T., Ishikawa K.-I., Sayama M., Kikuno R., Hiroshima M.,  
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
 RT Prediction of the coding sequences of unidentified human genes. XII.  
 The complete sequences of 100 new cDNA clones from brain which code  
 for large proteins in vitro.;  
 RL DNA Res. 5:355-364 (1998).  
 RN [3]  
 RP SEQUENCE OF 885-1972 FROM N.A.  
 RX MEDLINE=93263189; PubMed=7684189;  
 RA Matsuoka R., Yoshida M.C., Furutani Y., Imamura S., Kanda N.,  
 RA Yamagawa M., Masaki T., Takaoka A.;  
 RT "Human smooth muscle myosin heavy chain gene mapped to chromosomal  
 RT region 16q12.";  
 RL Am. J. Med. Genet. 46:61-67 (1993).  
 RN [4]  
 RP SEQUENCE OF 1093-1972 FROM N.A.  
 RC TISSUE=Hippocampus;  
 RA Okajima K.;  
 CC Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: MUSCLE CONTRACTION.  
 CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -!- TISSUE SPECIFICITY: SMOOTH MUSCLE; EXPRESSED IN THE UMBILICAL  
 CC ARTERY, BLADDER, ESOPHAGUS AND TRACHEA.  
 CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC INVERSION (INV16) (P13Q22), PRODUCES A FUSION PROTEIN THAT CONSISTS  
 CC OF THE 165 N-TERMINAL RESIDUES OF CBF-BETA (PEPB2) WITH THE TAIL  
 CC REGION OF MYH11. THIS RERARRGEMENT IS ASSOCIATED WITH ACUTE  
 CC MYELOID LEUKEMIA OF M4EO SUBTYPE.  
 CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MEROYOSIN (LMM) AND 1 HEAVY MEROYOSIN (HMM). IT CAN LATER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@sb-sib.ch).  
 CC  
 CC EMBL: AR001548; AAC31665.1; -.  
 DR EMBL: U91323; AAC35212.1; -.  
 DR EMBL: AB020673; BA74889.1; -.  
 DR EMBL: DQ0667; ; NOT\_ANNOTATED\_CDS.  
 DR EMBL: X99292; CAA9154.1; -.  
 DR HSSP: SMART; 1MNN.  
 DR MIM: 160745; -.  
 DR InterPro: IPRO00048; IQ.  
 DR InterPro: IPRO02938; Myosin\_tail.  
 DR InterPro: IPRO02017; Spectrin.  
 DR InterPro: IPRO01699; myosin\_head.  
 DR Pfam: PF00612; IQ\_1.  
 DR Pfam: PF0063; myosin\_head; 1.  
 DR Pfam: PF01576; Myosin\_tail; 1.  
 DR PRINTS: PRO0193; MYOSINHEAVY.  
 DR PRODOM: PD00335; myosin\_head; 1.  
 DR SMART; SM0015; IQ; 2.  
 DR SMART; SM0242; MSC; 1.  
 DR PROSITE; PS50096; IQ; 1.  
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;  
 KW Multigene family; Proto-oncogene; Chromosomal translocation.

FT DOMAIN 1  
 FT DOMAIN 786 815 MYOSIN HEAD-LIKE.  
 FT DOMAIN 844 1934 IQ.  
 FT DOMAIN 1935 1972 COILED COIL (POTENTIAL).  
 FT DOMAIN 178 185 CARBOXYL-TERMINAL.  
 FT NP-BIND 661 683 ATP (POTENTIAL).  
 FT DOMAIN 766 776 ACTIN-BINDING (BY SIMILARITY).  
 FT MOD-RES 129 129 METHYLATION (TRR-) (POTENTIAL).  
 FT MOD-RES 701 701 ALKYLATION (SH-1) (POTENTIAL).  
 FT MOD-RES 711 711 ALKYLATION (SH-2) (POTENTIAL).  
 FT CONFLICT 889 889 EEK -> NSE (IN REF. 3).  
 FT CONFLICT 1263 1266 ELQS -> TLSF (IN REF. 2).  
 FT CONFLICT 1558 1558 T -> S (IN REF. 3).  
 FT CONFLICT 1611 1611 KQ -> NE (IN REF. 3).  
 FT CONFLICT 1786 1786 A -> S (IN REF. 4).  
 FT CONFLICT 1958 1958 T -> L (IN REF. 3).  
 SQ SEQUENCE 1972 AA; 227338 MW; 67665BB2AEC1277 CRC64;

Query Match 10.7%; Score 95; DB 1; Length 1972;  
 Best Local Similarity 24.5%; Pred. No. 8; Gaps 1;  
 Matches 27; Conservative 18; Mismatches 55; Indels 10; Gaps 1;  
 ID MYH8\_RABBIT STANDARD; PRT; 1972 AA.

QY 53 CDRGLRRAVMECRNTHLQLQELTEAQKQFDVQAATCNHTWMLMASLD----- 103  
 Db 1268 CSDGGRARALNDKVHKQNEVSVGMNEAQKAIKAKOVALVASLSQLQDQELIQQE 1327

QY 104 -AEKAQGKQVEELEGETTNIKLDQSAEYERLRENQVSVRDAK 152  
 Db 1328 TQQLNVSTKLROLEEERNLSQDQDLEMKAONKLRIISTUNIQDSK 1377

RESULT 15  
 MYH8\_RABBIT STANDARD; PRT; 1972 AA.

ID MYH8\_RABBIT STANDARD; PRT; 1972 AA.

AC P31748; 01-JUN-1994 (Rel. 29, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Myosin heavy chain, smooth muscle isoform (SMMHC).  
 GN MYH11.

OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID:9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92073350; PubMed=1961735;  
 RA Babu J P., Kelly C., Periasamy M.;  
 RT Characterization of a mammalian smooth muscle myosin heavy-chain  
 gene: complete nucleotide and protein coding sequence and analysis of  
 RT the 5' end of the gene.;  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:10676-10680 (1991).

CC -!- FUNCTION: MUSCLE CONTRACTION.  
 CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MEROYOSIN (LMM) AND 1 HEAVY MEROYOSIN (HMM). IT CAN LATER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch).

DR	EMBL; M77812; RAA31395.1; -.
DR	PIR; A41604; A41604.
DR	HSSP; P08799; IMM3D.
DR	InterPro; IPR000048; IQ.
DR	InterPro; IPR004009; Myosin_N.
DR	InterPro; IPR002928; Myosin_tail.
DR	InterPro; IPR002017; Spectrin.
DR	InterPro; IPR001609; myosin_head.
DR	Pfam; PF00612; IQ; 1.
DR	Pfam; PF00063; myosin_head; 1.
DR	Pfam; PF02736; Myosin_N; 1.
DR	Pfam; PF01576; Myosin_tail; 1.
DR	PRINTS; PRO0133; MYOSINHEAVY.
DR	PRODOM; PD000355; myosin_head; 1.
DR	SMART; SM00155; IQ; 1.
DR	SMART; SM00242; MYC; 1.
PROSITE; PS50056; IO; 1.	Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding; Calmodulin-binding; ATP-binding; Methylation; Alkylation; Multigene family.
FT	DOMAIN 1 784 MYOSIN HEAD-LIKE.
FT	DOMAIN 1 785 807 IQ.
FT	DOMAIN 1 844 1934 COILED COIL (POTENTIAL).
FT	DOMAIN 1 1935 1972 CARBOXYL-TERMINAL.
FT	NP_01784 1 185 ATP (POTENTIAL).
FT	DOMAIN 1 651 683 ACTIN-BINDING (BY SIMILARITY).
FT	DOMAIN 1 763 777 ACTIN-BINDING (BY SIMILARITY).
FT	MOD_RES 1 129 129 METHYLATION (THI-1) (POTENTIAL).
FT	MOD_RES 1 701 701 ALKYLATION (SH-1) (POTENTIAL).
FT	MOD_RES 1 711 711 ALKYLIATION (SH-2) (POTENTIAL).
SEQUENCE	1972 AA; 227318 MW; 2061A224288BD6A4C CRC64;
Query	Match 10.7%; Score 95; DB 1; Length 1972;
Best Local Similarity 24.5%; Pred. No. 8;	
Matches 27; Conservative 18; Mismatches 55; Indels 10; Gaps 1;	
Qy 53	CRDGIRAVMCRNTHLQLQELTERQKGQDVEAQATCANTHTVMAIASID----- 103
Db 1268	CSGDGERARARALNDKVKHLQLNEVESVTGMLSEAEKGAIKLAKEVASLGSQQLDQIQLLQEE 1327
Qy 104	-AEKAQGOKKVEELEGEITITLNHKLQDASAEVERLRLRRENOVLSVRIADK 152
Db 1328	TRQKINVSTKLQJEDERINSLQEQDEEMAKONIHLRISTLNLQSDR 1377

Search completed: July 18, 2002, 18:44:50  
Job time: 128 sec

Job time: 128 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5  
copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 18:40:57 ; Search time 16.22 Seconds  
(without alignments)  
1066.343 Million cell updates/sec

Title: US-09-828-217-1  
Perfect score: 889  
Sequence: 1 MASTSYDCRVPMEDGDKRC . . . . . SSAAPOLLIVLVLGLSALLQ 180

Scoring table: BLOSUM62  
Gapext 10.0 , Gapext 0.5

Searched: 283138 seqs, 96080334 residues

Total number of hits satisfying chosen parameters: 283138

Maximum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 71.1:  
1: pir1:  
2: pir2:  
3: pir3:  
4: pir4:  
5: pir5:  
6: pir6:  
7: pir7:  
8: pir8:  
9: pir9:  
10: pir10:  
11: pir11:  
12: pir12:  
13: pir13:  
14: pir14:  
15: pir15:  
16: pir16:  
17: pir17:  
18: pir18:  
19: pir19:  
20: pir20:  
21: pir21:  
22: pir22:  
23: pir23:  
24: pir24:  
25: pir25:  
26: pir26:  
27: pir27:  
28: pir28:  
29: pir29:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	889	100.0	180	A56836
2	111	12.5	778	T30430
3	108	12.1	423	I55463
4	107.5	12.1	461	H84099
5	107	12.0	143	G13955
6	102.5	11.5	1534	A56734
7	102	11.5	2	A20166
8	99.5	11.2	304	JX0209
9	99	11.1	415	S35760
10	98	11.1	1963	MWKW
11	98	11.0	959	A55913
12	98	11.0	1938	MWKW
13	98	11.0	A59293	
14	97	10.9	286	S07533
15	96.5	10.9	597	S40998
16	96.5	10.9	893	2
17	96	10.8	286	S07532
18	95.5	10.7	173	S76705
19	95	10.7	1937	2
20	95	10.7	1972	1
21	94.5	10.6	1388	S74245
22	94	10.6	676	2
23	93.5	10.5	140	H64629
24	93.5	10.5	1938	1
25	93.5	10.5	1939	1
26	93	10.5	359	2
27	93	10.5	848	A44972
28	93	10.5	879	2
29	93	10.5	1938	2

## ALIGNMENTS

RESULT	1	2
A56836		
C;Species: Homo sapiens (man)		
C;Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 21-Jul-2000		
C;Accession: A56836		
R;Ishikawa, J.; Kishio, T.; Tomizawa, H.; Lee, B.O.; Kobune, Y.; Inazawa, J.; Oritani Genomics 26, 521-534, 1995		
A;Title: Molecular cloning and chromosomal mapping of a bone marrow stromal cell surf		
A;Reference number: A56836; MUID:95331788		
A;Accession: A56836		
A;Status: preliminary; not compared with conceptual translation		
A;Molecule type: mRNA		
A;Residues: 1-180 <TSD>		
A;Cross-references: GDB:DB8137; NIDN:BA05679.1; PID:9506861		
C;Genetics:		
A;Gene: GDB:BST2		
A;Cross-references: GDB:409946; OMIM:600534		
A;Map position: 19p13.2-19p13.2		
C;Keywords: transmembrane protein		
Query Match	100.0%	Score 889; DB 2; length 180;
Best local Similarity	100.0%	Pred. No. 2.6e-62; Mismatches 0; Indels 0; Gaps 0;
Matches	180;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1	MASTSYDCRVPMEDGDKRCRCKLIGILVLVLLITVILGVPLIFTIKANSAACRCDLIRAV 60
Db	1	MASTSYDCRVPMEDGDKRCRCKLIGILVLVLLITVILGVPLIFTIKANSAACRCDLIRAV 60
QY	61	MECRNVWHLQQLTEAQKGFDQEQAQATCNTWMLMASIDAERQAGQKRYVEELEG 120
Db	61	MECRNVWHLQQLTEAQKGFDQEQAQATCNTWMLMASIDAERQAGQKRYVEELEG 120
QY	121	TTLNHKLQDASAVERLRENQVLSYRADKYYKPSQDSSAAAPQLLVLIGLISALLQ 180
Db	121	TTLNHKLQDASAVERLRENQVLSYRADKYYKPSQDSSAAAPQLLVLIGLISALLQ 180
RESULT	2	
T30430		
hypothetical protein ORF82 - Lymantria dispar nuclear polyhedrosis virus		
C;Species: Lymantria dispar nuclear polyhedrosis virus, LdMNPV		
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999		
C;Accession: T30430		
R;Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; R Virology 253, 17-34, 1999		
A;Title: Sequence, analysis and synthesis of the genome of a baculovirus pathogenic for Lymantri		
A;Reference number: 220836; MUID:99124785		
A;Accession: T30430		
A;Status: preliminary; translated from GB/EMBL/DDBJ		
A;Molecule type: DNA		
smooth muscle myosin heavy chain		
secreted 45 kd pro		
general stress pro		
myosin heavy chain		
alpha cardiac myosin II heavy chain		
paramyosin - Caeno hypothetical prote		
beta-myosin heavy chain		
myosin heavy chain		
myosin heavy chain		
myosin beta heavy chain		
myosin alpha heavy chain		
myosin heavy chain		
centrosome associa		

A;Residues: 1-778 <KUZZ>  
A;Cross-references: EML:AF081810; PIDN: AAC70268.1

Query Match 12.5%; Score 111; DB 2; Length 778;  
Best Local Similarity 32.5%; Pred. No. 0.48; Mismatches 38; Conservative 22; Indels 10; Gaps 4;  
Matches 52; Conservative 28; Mismatches 56; Indels 50; Gaps 8;

QY 46 IKA  
N  
S  
E  
A  
C  
R  
D  
G  
L  
R  
A  
V  
M  
E  
C  
R  
N  
Y  
H  
L  
Q  
E  
T  
A  
Q  
K  
V  
D  
I  
G  
L  
R  
K  
V  
D  
T  
N  
T  
R  
L  
P  
E  
T  
E  
B  
A  
L  
K  
A  
E  
Q  
A  
D  
A  
L  
R  
L  
Q  
A  
E  
I  
D  
S  
L  
K  
R  
- - - A  
E  
S  
B  
A  
D  
L  
R  
N  
R  
V  
A  
Q  
L  
E  
A  
E  
A  
S  
575 GAA  
D  
L  
O  
N  
R  
I  
A  
R  
L  
B  
A  
E  
S  
G  
L  
T  
R  
R  
Q  
E  
S  
A  
E  
V  
L  
A  
L  
R  
D  
E  
R  
T  
A  
A  
S  
Q  
D  
S  
S  
161  
627

RESULT 3

QY 39463  
K  
a  
t  
r  
a  
t  
i  
n,  
t  
y  
p  
e  
1,  
c  
y  
t  
o  
s  
k  
e  
l  
e  
t  
a  
l  
-  
m  
o  
u  
s  
e  
N  
a  
l  
t  
e  
r  
n  
a  
m  
e  
s:  
e  
n  
d  
o  
B  
c  
y  
t  
o  
k  
e  
r  
a  
t  
i  
n  
D  
C;Species: Mus musculus (house mouse)

C;Accession: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 10-Dec-1999

C;Accession: 159463; A25621; A28428; JT0406

R;Alonso, A.; Weber, T.; Jorcano, J.L.

Rox's Arch. Dev. Biol. 196, 16-21, 1987

A;Title: Cloning and characterization of keratin D, a murine endodermal cytoskeletal protein

A;Reference number: 159463

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Residues: 1-423 <RES>

A;Cross-references: GB:M36376; NID:9198587; PIDN:AAA39373.1; PID:9293682

J;Singer, P.A.; Trevor, K.; Oshima, R.G.

A;Title: Molecular cloning and characterization of the endo B cytokeratin expressed in F

A;Reference number: A25621; MUID:86085876

A;Molecule type: mRNA

A;Accession: A25621

A;Title: Molecular cloning and characterization of the endo B cytokeratin expressed in F

A;Residues: 1-243; 'D', 245-252, 'A', 254-423 <SIN>

A;Cross-references: GB:ML1886; NID:9198620; PIDN:AAA39390.1; PID:9293685

R;Oshima, R.G.; Trevor, K.; Shevinsky, L.H.; Ryder, O.A.; Cecena, G.

Genes Dev. 21, 505-516, 1988

A;Title: Identification of the gene coding for the gene coding for the endo B murine cytokeratin and its met

A;Reference number: A28428; MUID:88255838

A;Accession: A28428

A;Molecule type: DNA

A;Residues: 1-132 <OSH>

A;Cross-references: GB:Y00217; NID:919842; PIDN:CAA68365.1; PID:950843

R;Ichinose, Y.; Morita, T.; Zhang, F.; Srinivasan, S.; Tondella, M.L.C.; Matsumoto, Gene 70, 85-95, 1988

A;Title: Nucleotide sequence and structure of the mouse cytokeratin endoB gene.

A;Reference number: JT0406; MUID:89196920

RESULT 5

Query Match 12.1%; Score 108; DB 2; Length 423;  
Best Local Similarity 28.0%; Pred. No. 0.44; Mismatches 52; Conservative 28; Mismatches 56; Indels 50; Gaps 8;

Query Match 12.0%; Score 107; DB 2; Length 143;  
Best Local Similarity 22.1%; Pred. No. 0.17; Mismatches 31; Conservative 34; Mismatches 49; Indels 26; Gaps 4;

QY 44 F  
T  
I  
K  
A  
N  
S  
E  
A  
C  
R  
D  
G  
L  
R  
A  
V  
M  
E  
C  
R  
N  
Y  
H  
L  
Q  
E  
T  
A  
Q  
K  
V  
D  
I  
G  
L  
R  
K  
V  
D  
T  
N  
T  
R  
L  
P  
E  
T  
E  
B  
A  
L  
K  
A  
E  
Q  
A  
D  
A  
L  
R  
L  
Q  
A  
E  
I  
D  
S  
L  
K  
R  
- - - A  
E  
S  
B  
A  
D  
L  
R  
N  
R  
V  
A  
Q  
L  
E  
A  
E  
A  
S  
575 GAA  
D  
L  
O  
N  
R  
I  
A  
R  
L  
B  
A  
E  
S  
G  
L  
T  
R  
R  
Q  
E  
S  
A  
E  
V  
L  
A  
L  
R  
D  
E  
R  
T  
A  
A  
S  
Q  
D  
S  
S  
161  
627

Db 217 L  
E  
A  
Q  
I  
A  
S  
S  
G  
L  
T  
V  
D  
A  
P  
K  
S  
Q  
D  
L  
S  
K  
I  
M  
A  
R  
Q  
E  
A  
L  
G  
Q  
R  
E  
E  
L  
D  
K  
Y  
W  
S  
Q  
I  
E  
E  
S  
T  
V  
T  
V  
777 T  
L  
N  
H  
K  
L  
Q  
D  
A  
S  
A  
E  
V  
L  
R  
R  
E  
N  
Q  
V  
L  
S  
V  
R  
I  
A  
E  
T  
T  
L  
E  
R  
L  
Q  
T  
P  
E  
I  
D  
S  
M  
K  
Q  
N  
I  
L  
E  
N  
S  
G  
D  
V  
E  
A  
R  
Y  
K  
A  
D  
O  
N  
Q  
V  
L  
336

QY 174 G  
L  
S  
A  
L  
L  
179

Db 337 H  
L  
E  
S  
E  
L  
342

RESULT 4

QY 84099  
H  
8  
4  
0  
9  
9

Db 217 L  
E  
A  
Q  
I  
A  
S  
S  
G  
L  
T  
V  
D  
A  
P  
K  
S  
Q  
D  
L  
S  
K  
I  
M  
A  
R  
Q  
E  
A  
L  
G  
Q  
R  
E  
E  
L  
D  
K  
Y  
W  
S  
Q  
I  
E  
E  
S  
T  
V  
T  
V  
777 T  
L  
N  
H  
K  
L  
Q  
D  
A  
S  
A  
E  
V  
L  
R  
R  
E  
N  
Q  
V  
L  
S  
V  
R  
I  
A  
E  
T  
T  
L  
E  
R  
L  
Q  
T  
P  
E  
I  
D  
S  
M  
K  
Q  
N  
I  
L  
E  
N  
S  
G  
D  
V  
E  
A  
R  
Y  
K  
A  
D  
O  
N  
Q  
V  
L  
336

QY 93 H  
V  
M  
A  
L  
M  
A  
S  
D  
E  
A  
K  
Q  
C  
K  
V  
E  
L  
G  
I  
T  
L  
N  
H  
K  
L  
Q  
D  
A  
S  
A  
E  
V  
L  
R  
R  
E  
N  
Q  
V  
L  
S  
V  
R  
I  
A  
E  
T  
T  
L  
E  
R  
L  
Q  
T  
P  
E  
I  
D  
S  
M  
K  
Q  
N  
I  
L  
E  
N  
S  
G  
D  
V  
E  
A  
R  
Y  
K  
A  
D  
O  
N  
Q  
V  
L  
336

Db 60 K  
E  
V  
E  
K  
E  
L  
G  
D  
I  
T  
A  
B  
I  
E  
R  
L  
D  
K  
E  
V  
E  
T  
S  
G  
K  
I  
O  
K  
R  
E  
E  
I  
F  
E  
V  
Q  
A  
B  
I  
E  
R  
L  
E  
R  
A  
E  
R  
118

QY 33 I  
T  
I  
V  
L  
G  
V  
P  
L  
I  
I  
T  
I  
K  
A  
N  
S  
E  
A  
C  
R  
D  
G  
L  
R  
A  
V  
M  
E  
C  
R  
N  
Y  
H  
L  
Q  
E  
T  
A  
Q  
K  
V  
D  
I  
G  
L  
R  
K  
V  
D  
T  
N  
T  
R  
L  
P  
E  
T  
E  
B  
A  
L  
K  
A  
E  
Q  
A  
D  
A  
L  
R  
L  
Q  
A  
E  
I  
D  
S  
L  
K  
R  
- - - A  
E  
S  
B  
A  
D  
L  
R  
N  
R  
V  
A  
Q  
L  
E  
A  
E  
A  
S  
575 GAA  
D  
L  
O  
N  
R  
I  
A  
R  
L  
B  
A  
E  
S  
G  
L  
T  
R  
R  
Q  
E  
S  
A  
E  
V  
L  
A  
L  
R  
D  
E  
R  
T  
A  
A  
S  
Q  
D  
S  
S  
161  
627

QY 23 I  
L  
I  
G  
I  
V  
L  
I  
I  
V  
L  
G  
V  
P  
L  
I  
I  
T  
I  
K  
A  
N  
S  
E  
A  
C  
R  
D  
G  
L  
R  
A  
V  
M  
E  
C  
R  
N  
Y  
H  
L  
Q  
E  
T  
A  
Q  
K  
V  
D  
I  
G  
L  
R  
K  
V  
D  
T  
N  
T  
R  
L  
P  
E  
T  
E  
B  
A  
L  
K  
A  
E  
Q  
A  
D  
A  
L  
R  
L  
Q  
A  
E  
I  
D  
S  
L  
K  
R  
- - - A  
E  
S  
B  
A  
D  
L  
R  
N  
R  
V  
A  
Q  
L  
E  
A  
E  
A  
S  
575 GAA  
D  
L  
O  
N  
R  
I  
A  
R  
L  
B  
A  
E  
S  
G  
L  
T  
R  
R  
Q  
E  
S  
A  
E  
V  
L  
A  
L  
R  
D  
E  
R  
T  
A  
A  
S  
Q  
D  
S  
S  
161  
627

Db 6 L  
Y  
N  
I  
M  
I  
L  
I  
I  
V  
L  
G  
V  
A  
V  
I  
F  
V  
N  
F  
N  
N  
E  
D  
Q  
R  
E  
P  
T  
I  
D  
I  
A  
S  
O  
V  
E  
T  
E  
T  
T  
N  
I  
L  
N  
D  
6

Qy	75	TER-----OKGFDQVEAQATCNHWMALMASLDAEKAQGKVKVELEGETTLNKK 126	Db	66	VRAFLIHVDNRNALQEVQKRDQFQVNNTIRSLAGMDASQLSGADGIEERLEAQ----- 118
RESULT	6				
A56734		ribosome receptor, 180k - dog			
C;Species:	Canis lupus familiaris (dog)				
C;Accession:	A56734	#sequence_revision 03-Aug-1995 #text_change 08-Oct-1999			
R;Wanker, E.E.; Sun, Y.; Savitz, A.J.; Meyer, D.I.					
J. Cell Biol. 130, 29-39, 1995					
A;Title:	Functional characterization of the 180-kD ribosome receptor in vivo.				
A;Reference number:	A56734; MUID:95310363				
A;Status:	preliminary; not compared with conceptual translation				
Molecule type: mRNA					
Residues: 1-1534 <WAN>					
A;Cross-references:	GB:X87224; NID:984113; PIDN:CAA60676.1; PID:984114				
C;Keywords:	endoplasmic reticulum; membrane protein; protein biosynthesis				
F;198-743/Region: 10-residue repeats (N-Q-G-K-K-A-E-G-A-P)					
Query Match	11.5%	Score 102.5; DB 2; Length 1534;			
Best Local Similarity	28.6%	Pred. No. 4.4; 42; Indels 1; Gaps 1;			
Matches	26;	Conservative 22; Mismatches			
A28166		Kupffer cell receptor - rat			
C;Species:	Rattus norvegicus (Norway rat)				
C;Date:	30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 05-Nov-1999				
C;Accession:	A38674; A28166				
R;Hoyle, G.W.; Hill, R.I.					
J. Biol. Chem. 266, 1850-1857, 1991					
A;Title:	Structure of the gene for a carbohydrate-binding receptor unique to rat Kupffer				
A;Reference number:	A38674; MUID:91107689				
A;Status:	preliminary				
A;Molecule type: DNA					
A;Residues: 1-550 <HO2>					
A;Cross-references:	GB:M55532; NID:9203362; PIDN:AAA40892.1; PID:9203363				
A;Molecule type: mRNA					
A;Accession:	A28166				
J. Biol. Chem. 263, 7487-7492, 1988					
A;Title:	Molecular cloning and sequencing of a cDNA for a carbohydrate binding receptor				
A;Reference number:	A28166; MUID:88227939				
Query Match	11.5%	Score 102; DB 2; Length 550;			
Best Local Similarity	28.0%	Pred. No. 1.7; 61; Indels 14; Gaps 3;			
Matches	37;	Conservative 20; Mismatches			
RESULT	7				
A28166		Kupffer cell receptor - rat			
C;Species:	Rattus norvegicus (Norway rat)				
C;Accession:	A38674; A28166				
R;Hoyle, G.W.; Hill, R.I.					
J. Biol. Chem. 266, 1850-1857, 1991					
A;Title:	Structure of the gene for a carbohydrate-binding receptor unique to rat Kupffer				
A;Reference number:	A38674; MUID:91107689				
A;Status:	preliminary				
A;Molecule type: DNA					
A;Residues: 1-550 <HO2>					
A;Cross-references:	GB:M55532; NID:9203362; PIDN:AAA40892.1; PID:9203363				
A;Molecule type: mRNA					
A;Accession:	A28166				
J. Biol. Chem. 263, 7487-7492, 1988					
A;Title:	Molecular cloning and sequencing of a cDNA for a carbohydrate binding receptor				
A;Reference number:	A28166; MUID:88227939				
Query Match	11.5%	Score 102; DB 2; Length 550;			
Best Local Similarity	28.0%	Pred. No. 1.7; 61; Indels 14; Gaps 3;			
Matches	37;	Conservative 20; Mismatches			
RESULT	8				
Oy	24	IGIGIIVLILIVLGVPLIFITKANSEACRQ---GLRAVMECRNVTHLQQLTEAQKGF 81	Db	41	IGLSLLILVWWSVIG-----SNSQLRRDILGTLRATLD--NTTSKIAE-----F 83
C;Species:	Streptococcus pyogenes				
C;Accession:	S35760; A42711				
R;Podbielski, A.					
R;Submitted to the EMBL Data Library, November 1992					
A;Reference number:	S35760				
A;Accession:	S35760				
A;Status: preliminary					
A;Molecule type: DNA					
A;Cross-references:	EMBL:X69324; NID:9311759; PIDN:CAA49165.1; PID:9311760				
R;Haase, F.T.; Heath, D.G.; Cleary, P.P.					
Qy	127	LODASAVERLRENQVLSW 146	Db	119	LQD--DINALMKGESVVKI 135
Qy	166	POLTVLIGLSA 177	Db	362	AEVQSLRTGLEA 373

J	Bacteriol	174	4967-4976	1992	A; Residues: 1873-1963 <W13>
A; Title:	Architecture of the vir regulons of group A streptococci parallels opacity fact			A; Cross-references: GB:V01494; GB:J01049; NID:96783; PIDN:CAA24738.1; PID:96784	
A; Reference number:	A42711; MUID:92332431			A; Gene: unc-54; CESP:F11C3.3	
A; Status:	preliminary			A; Map Position: 1	
A; Molecule type:	DNA			A; Introns: 21/3; 64/3; 111/3; 264/1; 525/3; 951/2; 1747/3; 1819/3; 1894/3	
A; Residues:	343-415 <HAA>			C; Superfamily: myosin heavy chain; myosin motor domain homology	
A; Cross-references:	GB:W6606; NID:9153630; PIDN:AAA26887.1; PID:9153631			C; Keywords: actin binding; AT; coiled coil; hydrolase; methylated amino acid; muscle	
A; Experimental source:	strain CS101; OF+			F; 184-175/Domain: myosin motor domain homology <MM01>	
A; Note:	sequence extracted from NCBI backbone (NCBIN:108942, NCBIPI:108945)			F; 174-181/Region: nucleotide-binding motif A (P-loop)	
A; Superfamily:	M5 protein			F; 662-680/Region: actin binding #status predicted	
A; Molecule type:	DNA			F; 766-780/Region: actin binding #status predicted	
A; Accession:	A21074			F; 848-1963/Domain: coiled coil #status predicted <<O1>>	
A; Cross-references:				F; 848-1162/Region: S2	
A; Reference number:				F; 1163-1963/Region: light meromyosin	
A; Status:	best local similarity			F; 125/Modified site: N6, N6-N6-trimethyllysine (Lys) #status predicted	
A; Molecule type:	DNA			F; 180/Binding site: ATP (Lys) #status predicted	
A; Accession:	A21074			F; 702-712/Active site: Cys #status predicted	
RESULT 10					
MRKK	myosin heavy chain B [similarity] - Caenorhabditis elegans	11.1%	Score 99; DB 2; Length 415;	Query Match	
N; Contains:	myosin ATPase (EC 3.6.1.32)	26.9%	Pred. No. 2.2; Mismatches 35; Conservative 29; Indels 20; Gaps 5;	Best Local Similarity	
C; Species:	Caenorhabditis elegans	35	Matches 35; Conservative 29; Mismatches 46; Indels 20; Gaps 5;	Conservative	
C; Date:	13-Jun-1983 #sequence_revision 19-May-2000 #text_change 19-Jan-2001			Indels 31; Gaps 5;	
C; Accession:	T20770; T21629; A93958; A93287; A21074; A02992			Matches 36; Conservative 21; Mismatches 58; Indels 31; Gaps 5;	
R; Kershaw, J. submitted to the EMBL Data Library, November 1996					
A; Reference number:	Z19322				
A; Status:	translated from GB/EMBL/DBJ				
A; Molecule type:	DNA				
A; Residues:	1-1963 <W13>				
A; Cross-references:	EMBL:Z81499; PIDN:CA04089.1; GSFD:GN00019; CESP:F11C3.3				
A; Experimental source:	clone FL1C3				
A; Accession:	T21629				
A; Status:	translated from GB/EMBL/DBJ				
A; Molecule type:	DNA				
A; Residues:	1-1963 <W12>				
A; Cross-references:	EMBL:Z83107; PIDN:CA05505.1; GSFD:GN00019; CESP:F11C3.3				
A; Experimental source:	clone F3A7				
R; Kurn, J.; Brenner, S.; Barnett, L.					
Proc. Natl. Acad. Sci. U.S.A. 80, 4233-4257, 1983					
A; Title:	Protein structural domains in the Caenorhabditis elegans unc-54 myosin heavy chain				
A; Reference number:	A93958; MUID:83273600				
A; Accession:	A93958				
A; Molecule type:	DNA				
A; Residues:	1-61, EMSVQ'-65-376, 'V' 378-1963 <KAR>				
A; Cross-references:	GB:J01050; NID:9153639; PIDN:AA28124.1; PID:9156400				
A; Title:	Periodic charge distributions in the myosin rod amino acid sequence match cross				
A; Reference number:	A93287; MUID:82272395				
A; Accession:	A93287				
A; Molecule type:	DNA				
A; Residues:	847-1333, 'R' 1335-1876, 'L' 1878-1963 <W13>				
R; Wills, N.; Gestrelle, R.F.; Kurn, J.; Barnett, L.; Bolten, S.; Waterston, R.H.					
Cell 33, 575-583, 1983					
A; Title:	The genes sup-7 X and sup-5 III of Caenorhabditis elegans suppress amber nonsense				
A; Reference number:	A21074; MUID:83232892				
A; Accession:	A21074				
A; Molecule type:	DNA				
RESULT 11					
MRKK	myosin heavy chain B [similarity] - Caenorhabditis elegans	11.1%	Score 99; DB 2; Length 415;	Query Match	
N; Contains:	myosin ATPase (EC 3.6.1.32)	26.9%	Pred. No. 2.2; Mismatches 35; Conservative 29; Indels 20; Gaps 5;	Best Local Similarity	
C; Species:	Caenorhabditis elegans	35	Matches 35; Conservative 29; Mismatches 46; Indels 20; Gaps 5;	Conservative	
C; Date:	13-Jun-1983 #sequence_revision 19-May-2000 #text_change 19-Jan-2001			Indels 31; Gaps 5;	
C; Accession:	T20770; T21629; A93958; A93287; A21074; A02992			Matches 36; Conservative 21; Mismatches 58; Indels 31; Gaps 5;	
R; Kershaw, J. submitted to the EMBL Data Library, November 1996					
A; Reference number:	Z19322				
A; Status:	translated from GB/EMBL/DBJ				
A; Molecule type:	DNA				
A; Residues:	1-1963 <W13>				
A; Cross-references:	EMBL:Z81499; PIDN:CA04089.1; GSFD:GN00019; CESP:F11C3.3				
A; Experimental source:	clone FL1C3				
A; Accession:	T21629				
A; Status:	translated from GB/EMBL/DBJ				
A; Molecule type:	DNA				
A; Residues:	1-1963 <W12>				
A; Cross-references:	EMBL:Z83107; PIDN:CA05505.1; GSFD:GN00019; CESP:F11C3.3				
A; Experimental source:	clone F3A7				
R; Kurn, J.; Barnett, S.; Barnett, L.					
Proc. Natl. Acad. Sci. U.S.A. 80, 4233-4257, 1983					
A; Title:	Protein structural domains in the Caenorhabditis elegans unc-54 myosin heavy chain				
A; Reference number:	A93958; MUID:83273600				
A; Accession:	A93958				
A; Molecule type:	DNA				
A; Residues:	1-61, EMSVQ'-65-376, 'V' 378-1963 <KAR>				
A; Cross-references:	GB:J01050; NID:9153639; PIDN:AA28124.1; PID:9156400				
A; Title:	Periodic charge distributions in the myosin rod amino acid sequence match cross				
A; Reference number:	A93287; MUID:82272395				
A; Accession:	A93287				
A; Molecule type:	DNA				
A; Residues:	847-1333, 'R' 1335-1876, 'L' 1878-1963 <W13>				
R; Wills, N.; Gestrelle, R.F.; Kurn, J.; Barnett, L.; Bolten, S.; Waterston, R.H.					
Cell 33, 575-583, 1983					
A; Title:	The genes sup-7 X and sup-5 III of Caenorhabditis elegans suppress amber nonsense				
A; Reference number:	A21074; MUID:83232892				
A; Accession:	A21074				
A; Molecule type:	DNA				
RESULT 12					
MRKK	myosin heavy chain B [similarity] - Caenorhabditis elegans	11.1%	Score 99; DB 2; Length 415;	Query Match	
N; Contains:	myosin ATPase (EC 3.6.1.32)	26.9%	Pred. No. 2.2; Mismatches 35; Conservative 29; Indels 20; Gaps 5;	Best Local Similarity	
C; Species:	Caenorhabditis elegans	35	Matches 35; Conservative 29; Mismatches 46; Indels 20; Gaps 5;	Conservative	
C; Date:	13-Jun-1983 #sequence_revision 19-May-2000 #text_change 19-Jan-2001			Indels 31; Gaps 5;	
C; Accession:	T20770; T21629; A93958; A93287; A21074; A02992			Matches 36; Conservative 21; Mismatches 58; Indels 31; Gaps 5;	
R; Kershaw, J. submitted to the EMBL Data Library, November 1996					
A; Reference number:	Z19322				
A; Status:	translated from GB/EMBL/DBJ				
A; Molecule type:	DNA				
A; Residues:	1-1963 <W13>				
A; Cross-references:	EMBL:Z81499; PIDN:CA04089.1; GSFD:GN00019; CESP:F11C3.3				
A; Experimental source:	clone FL1C3				
A; Accession:	T21629				
A; Status:	translated from GB/EMBL/DBJ				
A; Molecule type:	DNA				
A; Residues:	1-1963 <W12>				
A; Cross-references:	EMBL:Z83107; PIDN:CA05505.1; GSFD:GN00019; CESP:F11C3.3				
A; Experimental source:	clone F3A7				
R; Kurn, J.; Barnett, S.; Barnett, L.					
Proc. Natl. Acad. Sci. U.S.A. 80, 4233-4257, 1983					
A; Title:	Protein structural domains in the Caenorhabditis elegans unc-54 myosin heavy chain				
A; Reference number:	A93958; MUID:83273600				
A; Accession:	A93958				
A; Molecule type:	DNA				
A; Residues:	1-61, EMSVQ'-65-376, 'V' 378-1963 <KAR>				
A; Cross-references:	GB:J01050; NID:9153639; PIDN:AA28124.1; PID:9156400				
A; Title:	Periodic charge distributions in the myosin rod amino acid sequence match cross				
A; Reference number:	A93287; MUID:82272395				
A; Accession:	A93287				
A; Molecule type:	DNA				
A; Residues:	847-1333, 'R' 1335-1876, 'L' 1878-1963 <W13>				
R; Wills, N.; Gestrelle, R.F.; Kurn, J.; Barnett, L.; Bolten, S.; Waterston, R.H.					
Cell 33, 575-583, 1983					
A; Title:	The genes sup-7 X and sup-5 III of Caenorhabditis elegans suppress amber nonsense				
A; Reference number:	A21074; MUID:83232892				
A; Accession:	A21074				
A; Molecule type:	DNA				
RESULT 13					
MRKK	myosin heavy chain B [similarity] - Caenorhabditis elegans	11.1%	Score 99; DB 2; Length 415;	Query Match	
N; Contains:	myosin ATPase (EC 3.6.1.32)	26.9%	Pred. No. 2.2; Mismatches 35; Conservative 29; Indels 20; Gaps 5;	Best Local Similarity	
C; Species:	Caenorhabditis elegans	35	Matches 35; Conservative 29; Mismatches 46; Indels 20; Gaps 5;	Conservative	
C; Date:	13-Jun-1983 #sequence_revision 19-May-2000 #text_change 19-Jan-2001			Indels 31; Gaps 5;	
C; Accession:	T20770; T21629; A93958; A93287; A21074; A02992			Matches 36; Conservative 21; Mismatches 58; Indels 31; Gaps 5;	
R; Kershaw, J. submitted to the EMBL Data Library, November 1996					
A; Reference number:	Z19322				
A; Status:	translated from GB/EMBL/DBJ				
A; Molecule type:	DNA				
A; Residues:	1-1963 <W13>				
A; Cross-references:	EMBL:Z81499; PIDN:CA04089.1; GSFD:GN00019; CESP:F11C3.3				
A; Experimental source:	clone FL1C3				
A; Accession:	T21629				
A; Status:	translated from GB/EMBL/DBJ				
A; Molecule type:	DNA				
A; Residues:	1-1963 <W12>				
A; Cross-references:	EMBL:Z83107; PIDN:CA05505.1; GSFD:GN00019; CESP:F11C3.3				
A; Experimental source:	clone F3A7				
R; Kurn, J.; Barnett, S.; Barnett, L.					
Proc. Natl. Acad. Sci. U.S.A. 80, 4233-4257, 1983					
A; Title:	Protein structural domains in the Caenorhabditis elegans unc-54 myosin heavy chain				
A; Reference number:	A93958; MUID:83273600				
A; Accession:	A93958				
A; Molecule type:	DNA				
A; Residues:	1-61, EMSVQ'-65-376, 'V' 378-1963 <KAR>				
A; Cross-references:	GB:J01050; NID:9153639; PIDN:AA28124.1; PID:9156400				
A; Title:	Periodic charge distributions in the myosin rod amino acid sequence match cross				
A; Reference number:	A93287; MUID:82272395				
A; Accession:	A93287				
A; Molecule type:	DNA				
A; Residues:	847-1333, 'R' 1335-1876, 'L' 1878-1963 <W13>				
R; Wills, N.; Gestrelle, R.F.; Kurn, J.; Barnett, L.; Bolten, S.; Waterston, R.H.					
Cell 33, 575-583, 1983					
A; Title:	The genes sup-7 X and sup-5 III of Caenorhabditis elegans suppress amber nonsense				
A; Reference number:	A21074; MUID:83232892				
A; Accession:	A21074				
A; Molecule type:	DNA				
RESULT 14					
MRKK	myosin heavy chain B [similarity] - Caenorhabditis elegans	11.1%	Score 99; DB 2; Length 415;	Query Match	
N; Contains:	myosin ATPase (EC 3.6.1.32)	26.9%	Pred. No. 2.2; Mismatches 35; Conservative 29; Indels 20; Gaps 5;	Best Local Similarity	
C; Species:	Caenorhabditis elegans	35	Matches 35; Conservative 29; Mismatches 46; Indels 20; Gaps 5;	Conservative	
C; Date:	13-Jun-1983 #sequence_revision 19-May-2000 #text_change 19-Jan-2001			Indels 31; Gaps 5;	
C; Accession:	T20770; T21629; A93958; A93287; A21074; A02992			Matches 36; Conservative 21; Mismatches 58; Indels 31; Gaps 5;	
R; Kershaw, J. submitted to the EMBL Data Library, November 1996					
A; Reference number:	Z19322				
A; Status:	translated from GB/EMBL/DBJ				
A; Molecule type:	DNA				
A; Residues:	1-1963 <W13>				
A; Cross-references:	EMBL:Z81499; PIDN:CA04089.1; GSFD:GN00019; CESP:F11C3.3				
A; Experimental source:	clone FL1C3				
A; Accession:	T21629				
A; Status:	translated from GB/EMBL/DBJ				
A; Molecule type:	DNA				
A; Residues:	1-1963 <W12>				
A; Cross-references:	EMBL:Z83107; PIDN:CA05505.1; GSFD:GN00019; CESP:F11C3.3				
A; Experimental source:	clone F3A7				
R; Kurn, J.; Barnett, S.; Barnett, L.					
Proc. Natl. Acad. Sci. U.S.A. 80, 4233-4257, 1983					
A; Title:	Protein structural domains in the Caenorhabditis elegans unc-54 myosin heavy chain				
A; Reference number:	A93958; MUID:83273600				
A; Accession:	A93958				
A; Molecule type:	DNA				
A; Residues:	1-61, EMSVQ'-65-376, 'V' 378-1963 <KAR>				
A; Cross-references:	GB:J01050; NID:9153639; PIDN:AA28124.1; PID:9156400				
A; Title:	Periodic charge distributions in the myosin rod amino acid sequence match cross				
A; Reference number:	A93287; MUID:82272395				
A; Accession:	A93287				
A; Molecule type:	DNA				
A; Residues:	847-1333, 'R' 1335-1876, 'L' 1878-1963 <W13>				
R; Wills, N.; Gestrelle, R.F.; Kurn, J.; Barnett, L.; Bolten, S.; Waterston, R.H.					
Cell 33, 575-583, 1983					
A; Title:	The genes sup-7 X and sup-5 III of Caenorhabditis elegans suppress amber nonsense				
A; Reference number:	A21074; MUID:83232892				
A; Accession:	A21074				
A; Molecule type:	DNA				

myosin heavy chain D [similarity] - caenorhabditis elegans  
 N: Alternate names: myosin heavy chain I  
 N: Contains: myosin ATPase (BC 3.6.1.32)  
 C: Species: Caenorhabditis elegans  
 C: Date: 28-Feb-1986 #sequence\_revision 19-May-2000 #text\_change 19-Jan-2001  
 C: Accession: T21193; T23973; S02772; A02993  
 R: McMurray, A.  
 submitted to the EMBL Data Library, April 1996  
 A: Accession: T21193  
 A: Status: translated from GB/EMBL/DDBJ  
 A: Molecule type: DNA  
 A: Residues: 1-1938 <WIN>  
 A: Cross-references: EMBL:Z71261; PIDN:CAA95848.1; GSPDB:GN00019; CESP:R06C7.10  
 A: Experimental source: clone F21C3  
 A: Gardner, A.  
 submitted to the EMBL Data Library, April 1996  
 A: Reference number: Z19825  
 A: Accession: T23973  
 A: Status: translated from GB/EMBL/DDBJ  
 A: Molecule type: DNA  
 A: Residues: 1-1938 <WIN>  
 A: Cross-references: EMBL:Z71266; PIDN:CAA95848.1; GSPDB:GN00019; CESP:R06C7.10  
 A: Experimental source: clone R06C7  
 A: Dibb, N.J.; Matuyama, I.N.; Krause, M.; Karn, J.  
 J. Mol. Biol. 205, 603-613, 1989  
 A: Title: Sequence analysis of the complete *Caenorhabditis elegans* myosin heavy chain gene  
 A: Reference number: S02771; MUID:89178677  
 A: Status: nucleic acid sequence not shown  
 A: Molecule type: DNA  
 A: Residues: 1-376; 'V', 378-390; 'V', 392-576; 'L', 578-680; 'I', 682-1938 <DIB>  
 A: Cross-references: EMBL:X00665; NID:96785; PIDN:CAA30854.1; PID:96786  
 R: Kain, J.; Brenner, S.; Barnett, L.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 4253-4257, 1983  
 A: Title: Protein structural domains in the *Caenorhabditis elegans* unc-54 myosin heavy chain  
 A: Reference number: A93958; MUID:83273600  
 A: Accession: A02993  
 A: Molecule type: DNA  
 A: Residues: 24-93, 'E', 95-97, 'R', 99-376, 'V', 378-388, 'GDV', 392-407, 'N', 409-473, 'G', 475-576  
 C: Genetics:  
 A: Gene: myo-1, CESP:R06C7.10  
 A: Map position: 1  
 A: Introns: 23/3; 114/3; 229/1; 264/1; 320/1; 857/3; 1745/3; 1814/1; 1892/3  
 C: Superfamily: myosin heavy chain; myosin motor domain homology  
 C: Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; muscle cc  
 F: 87-773/Domain: myosin motor domain homology <MMOT>  
 F: 177-184/Region: nucleotide-binding motif A (P-1-loop)  
 F: 660-692/Region: actin binding #status predicted  
 F: 764-778/Region: actin binding #status predicted  
 F: 846-1938/Domain: coiled coil #status predicted <COI>  
 F: 846-1166/Region: S2  
 F: 1161-1938/Region: light meromyosin  
 F: 128/Modified site: N6, N6, N6-trimethyllysine (lys) #status predicted  
 F: 183/Binding site: ATP (lys) #status predicted  
 F: 700, 710/Active site: Cys #status predicted  
 RESULT 13  
 S09203  
 skeletal myosin heavy chain - domestic rabbit  
 C: Species: Oryctolagus cuniculus  
 C: Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 08-Sep-2000  
 C: Accession: A59293  
 R: Mada, K.; Hostinova, E.; Roess, Kleinkauf, A.; Schuster, H.; Gasperik, J.; Wittingh  
 submitted to GenBank, July 1995  
 A: Description: Isolation, sequencing of myosin heavy chain cDNA from rabbit skeletal  
 A: Reference number: A59293  
 A: Accession: A59293  
 A: Status: preliminary; not compared with conceptual translation  
 A: Molecule type: mRNA  
 A: Residues: 1-1938 <MAE>  
 A: Cross-references: GB:U32574; NID:940232; PIDN:AAA74199.1; PID:940233  
 A: Experimental source: strain New Zealand White; cell type skeletal muscle fiber type  
 C: Genetics:  
 A: Gene: MHC  
 C: Superfamily: myosin heavy chain; myosin motor domain homology <MMO>  
 F: 85-769/Domain: myosin motor domain homology  
 Query Match 11.0%; Score 98; DB 2; Length 1938;  
 Best Local Similarity 26.1%; Pred. No. 13; Indels 20; Gaps 5;  
 Matches 37; Conservative 20; Mismatches 45; Indels 40; Gaps 5;  
 QY 47 KANSKACR-----DGLRAVMECRVTHLQELTEAQKGFDQYEAQATCNTVMAL 99  
 Db 1366 KANSEVAQWRTKYEIDAIQTEELPEAKKLAQRLQDAE---EHVAEVAKC----- 1414  
 QY 100 ASLDLEKAQOKQKVPELEGTEITLN-----HKLQDASAVERLR 138  
 Db 1415 ASLETKRQLQNEVEDLMDIVERTVNAACAALDKKQRNFDFKILLAENKHKYBETHAILEASQ 1474  
 QY 139 RENQVSVRAD-KYVPPSSD 159  
 Db 1475 KESRSILSTEVFVKVKAAYEESLD 1496  
 RESULT 14  
 S07533  
 puff II/9A-2 protein precursor - fungus gnat (Sciara coprophila)  
 C: Species: Sciara coprophila  
 C: Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Jun-2000  
 C: Accession: S07533  
 R: Di Bartolomei, S.M.; Gerbi, S.A.  
 J. Mol. Biol. 210, 531-540, 1989  
 A: Title: Molecular characterization of DNA puff II/9A genes in *Sciara coprophila*.  
 A: Reference number: S07532; MUID:90133907  
 A: Accession: S07533  
 A: Status: not compared with conceptual translation  
 A: Molecule type: DNA  
 A: Residues: 1-286 <DIB>  
 A: Cross-references: GB:X151679; NID:910113; PID:91405812  
 C: Genetics:  
 A: Map position: II/9A  
 A:保守氨基酸: 24; 错配氨基酸: 31; 缺失氨基酸: 4; 插入氨基酸: 1  
 C: Map position: II/9A  
 C:保守氨基酸: 24; 错配氨基酸: 36; 缺失氨基酸: 30; 插入氨基酸: 5;  
 F: 156/Binding site: signal sequence #status predicted <SIG>  
 F: 21-286/Product: puff II/9A protein #status predicted <MAT>  
 F: 1-20/Domain: signal sequence #status predicted (covalent) #status predicted  
 Query Match 10.9%; Score 97; DB 2; Length 286;  
 Best Local Similarity 25.6%; Pred. No. 2,1; Indels 30; Gaps 5;  
 Matches 31; Conservative 24; Mismatches 36; Indels 30; Gaps 5;  
 QY 46 IKANSKACRCDGLRAVMEC-RNVTL-----LQQLTEAQKGFDQYEAQATCNTVMAL 98  
 Db 89 LKREKARQKEKAKLKECQKVNENKETIEQLKKELAQAOKALECKKELADCK----- 142  
 QY 164 AAPOLLIV 171  
 QY 1384 GEGLYGSEELEELKKRQMQNRMVMDLQEALSSA-----QNKVISLEKAKGKLLAETDARS 1438

Page 6

QY	99	MASLDAEKAQGOKKVERLEEGITTLNHKLQ-----DASAVERLRE-----NOQLSV	146
Db	143	-----KENAKLNLKIEBLNCTITQLOEKLERGRGRERDLQCQOLDECKKKLNITCNELIAC	197
Qy	147	R 147	
Db	198	R 198	

RESULT 15

540998

#### Hypothetical pr-

C;Species: *Caenorhabditis elegans* C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #te

C; Accession: S40998

R; Hawkins, T.; Thomas, R.  
submitted to the EMR Data Library October 1993

Submitted to the LMB Data Library, October 1993

Accession: S40998

Status: preliminary

A: Molecule Type: DNA  
A: Residues: 1-597 <HAW>

## A;Cross-references: EM

C; Genetics: 10/3: 25/1 105/1 3

A; millions: 40/3; 18/1; 125/1; 1

Best Local Similarity	Score	DB	Length
10.95%	90.5%	2	59;
32; Conservative	24.8%; Pred No. 4.9;	Mismatches	39; Indels 29; Gaps 4

40 INDEPENDENT INVESTIGATOR 101

BB 270 SNOQVIRNHANAV-ESLQKTHETTQIAEKNKEFFERNFEEERARREAECMAMNRHOKVAC 328

卷之三

102 LDAEKAGQOKKVEEL-----GEITLNHKQDASAEVERLRR 139

329 LDEK I SEAKO CEOLN VDKVLA ALAND CDH RHO MLTKEI SLOTA LEMSKA EMKEL RQ 388

140 ENZYME CIRCUIT 142

bb 389 KNONLSLOV 397

Search completed: July 18, 2002, 18:43:58  
b time: 181 sec

111

TM protein - protein search, using sw model  
 run on: July 18, 2002, 18:38:22 ; Search time 12.96 Seconds  
 (without alignments)  
 339.244 Million cell updates/sec

Copyright (c) 1993 - 2000 Compugen Ltd.

GenCore version 4.5

title: US-09-828-217-1  
 perfect score: 889  
 sequence: 1 MASTSYVCRVPMEDGDKRC. .... SSAAPQQLIVLGLSALQ 180  
 scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

searched: 231628 seqs, 2442594 residues

total number of hits satisfying chosen parameters: 231628  
 minimum DB seq length: 0  
 maximum DB seq length: 2000000000

post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

database : issued\_patents\_aa:\*

1: /ragn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*

2: /ragn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*

3: /ragn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*

4: /ragn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*

5: /ragn2\_6/ptodata/2/iaa/PCUTS\_COMB.pep:\*

6: /ragn2\_6/ptodata/2/iaa/backfilesl.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Query Length	DB ID	Description
1	889	100	180	2 US-08-624-650-1 Sequence 1, Appli
2	99.5	110	270	2 US-09-05-095-4 Sequence 4, Appli
3	99.5	112	270	2 US-08-899-49A-2 Sequence 2, Appli
4	99.5	112	270	4 US-09-352-302-2 Sequence 2, Appli
5	99.5	112	273	2 US-08-899-49A-2 Sequence 4, Appli
6	99.5	112	273	4 US-09-352-302-4 Sequence 4, Appli
7	95	107	285	2 US-08-533-300A-4 Sequence 4, Appli
8	95	107	285	2 US-08-742-92A-4 Sequence 4, Appli
9	93.5	105	1886	4 US-08-988-105-3 Sequence 3, Appli
10	92.5	104	292	2 US-08-608-342-4 Sequence 4, Appli
11	92.5	104	292	2 US-09-113-788-4 Sequence 4, Appli
12	89.5	101	316	4 US-09-111-470-4 Sequence 4, Appli
13	89.5	101	1939	4 US-09-330-18A-1 Sequence 1, Appli
14	88	99	288	3 US-08-312-949-4 Sequence 4, Appli
15	88	99	288	3 US-08-446-201-4 Sequence 4, Appli
16	88	99	619	1 US-08-455-746-2 Sequence 2, Appli
17	88	99	619	1 US-08-214-16A-2 Sequence 2, Appli
18	88	99	619	2 US-08-457-852A-3 Sequence 3, Appli
19	88	99	619	2 US-08-247-636-2 Sequence 2, Appli
20	88	99	619	2 US-08-247-491A-3 Sequence 3, Appli
21	88	99	619	2 US-08-319-795-2 Sequence 2, Appli
22	88	99	619	2 US-08-458-985-2 Sequence 2, Appli
23	88	99	619	3 US-08-312-949-2 Sequence 2, Appli
24	88	99	648	1 US-08-072-070-2 Sequence 2, Appli
25	88	99	648	1 US-08-439-434-2 Sequence 2, Appli
26	88	99	648	1 US-08-214-222-2 Sequence 2, Appli
27	88	99	648	2 US-08-457-852A-2 Sequence 2, Appli

RESULT	1	US-08-634-650-1
		Sequence 1, Application US/08624650
		Patent No. 5914252
		GENERAL INFORMATION:
		APPLICANT: HIRANO, TOSHIRO
		APPLICANT: KAI SHO, TSUNEYASU
		TITLE OF INVENTION: MEMBRANE PROTEIN POLYPEPTIDE HAVING PRE-B CELL GROWTH-SUPPORTING ABILITY AND A GENE THEREOF
		NUMBER OF SEQUENCES: 2
		CORRESPONDENCE ADDRESS:
		ADDRESSEE: OBLON, SRIKAK, MCCLELLAND, MAIER & NEUSTADT,
		STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
		CITY: ARLINGTON
		STATE: VA
		COUNTY: USA
		ZIP: 22202
		COMPUTER READABLE FORM:
		MEDIUM TYPE: FLOPPY disk
		COMPUTER: IBM PC compatible
		OPERATING SYSTEM: PC-DOS/MS-DOS
		SOFTWARE: PatentIn Release #1.0, Version #1.30
		CURRENT APPLICATION DATA:
		APPLICATION NUMBER: US/08/624,650
		APPLICATION NUMBER: JP 5-281622
		FILING DATE: 22-MAY-1996
		CLASSIFICATION: 435
		PRIOR APPLICATION DATA:
		APPLICATION NUMBER: PCT/JP94/01732
		FILING DATE: 14-OCT-1994
		PRIOR APPLICATION DATA:
		APPLICATION NUMBER: JP 5-281622
		FILED DATE: 15-OCT-1993
		ATTORNEY/AGENT INFORMATION:
		NAME: OBLON, NORMAN F.
		REGISTRATION NUMBER: 24,618
		REFERENCE/POCKET NUMBER: 7625-001-0 PCT
		TELECOMMUNICATION INFORMATION:
		TELEPHONE: 703-413-3000
		TELEFAX: 703-413-2220
		INFORMATION FOR SEQ ID NO: 1:
		SEQUENCE CHARACTERISTICS:
		LENGTH: 180 amino acids
		TYPE: amino acid
		STRANDEDNESS: single
		TOPOLOGY: linear
		MOLECULE TYPE: peptide
		US-08-624-650-1

Query Match 100.0%; Score 889; DB 2; Length 180;  
 Best Local Similarity 4.3%; Pred. No. 4; 3e-88;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDYCVPMECDKRKLLGIGIWLILITVIGVPLIFTIKANSEACRDGRAY 60  
 Db 1 MASTSYDYCVPMECDKRKLLGIGIWLILITVIGVPLIFTIKANSEACRDGRAY 60

QY 61 MECRNVTHLQQELTEAQKGFQDVEQAATCNHTWALMASLDAEKAGQKVEELEGET 120  
 Db 61 MECRNVTHLQQELTEAQKGFQDVEQAATCNHTWALMASLDAEKAGQKVEELEGET 120

QY 121 TTLNHLKLODASAEVERLRRNENQVLSVRIAKKYVSSQDSSAAQOLLVYLGLSALQ 180  
 Db 121 TTLNHLKLODASAEVERLRRNENQVLSVRIAKKYVSSQDSSAAQOLLVYLGLSALQ 180

RESULT 2  
 Sequence 4, Application US/09055095  
 Patent No. 5945308

GENERAL INFORMATION:  
 APPLICANT: Tang, Y. Tom  
 APPLICANT: Patterson, Chandra  
 APPLICANT: Coxley, Neil C.  
 APPLICANT: Sather, Susan

NUMBER OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR  
 NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Dr.  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304

COMPUTER READABLE FORM:  
 COMPUTER TYPE: Disquette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSEo for Windows Version 2.0

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/055,095  
 FILING DATE: Filed Herewith  
 CLASSIFICATION:  
 PRIORITY APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Billings, Lucy J.  
 REFERENCE/DOCKET NUMBER: 36,749  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650-855-0555  
 TELEFAX: 650-845-4166  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 270 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 LIBRARY: Geobank  
 CLONE: 1902982  
 US-09-055-095-4

Query Match 11.2%; Score 99.5%; DB 2; Length 270;  
 Best Local Similarity 25.9%; Pred. No. 0.0068; 41; Indels 37; Gaps 5;  
 Matches 36; Conservative 25; Mismatches 41; Indels 37; Gaps 5;

QY 23 LIGIGILVLLITVIGVPLIFTIKANSEACRDGRAYMCRNVTLLQOELTEAQKGFQ 82  
 Db 40 VLCIGLVLVWILLLQLSQVSDLIKQQ-----ANITH--QEDILEGO--- 80

QY 83 DVEQAATCNHTWALMASLDAEKAGQKVEELEGETTTLNHLKLODASAEVERLRRNQ 142  
 Db 81 -----ILAQRSSEKS-AQESQKELEMETLAHKDESKKLMELHRONL 124

RESULT 3  
 Sequence 2, Application US/08809494A  
 Patent No. 5962560

GENERAL INFORMATION:  
 APPLICANT: Sawamura, Tatsuya  
 APPLICANT: Masaki, Tomoo  
 TITLE OF INVENTION: Modified Low-Density Lipoprotein  
 NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: McAulay Fisher Nissen Goldberg & Kiel  
 STREET: 261 Madison Avenue  
 CITY: New York  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10016-2391

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/0809,494A  
 FILING DATE: 24-MAR-1997  
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 6-321705  
 FILING DATE: 30-NOV-1994

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 7-214206  
 FILING DATE: 31-JUL-1995

ATTORNEY/AGENT INFORMATION:  
 NAME: Goldberg, Jules E  
 REFERENCE/DOCKET NUMBER: 24408  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212 986-4090  
 TELEFAX: 212 818-9479

INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 20 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-809-494A-2

Query Match 11.2%; Score 99.5%; DB 2; Length 270;  
 Best Local Similarity 25.9%; Pred. No. 0.0068; 41; Indels 37; Gaps 5;  
 Matches 36; Conservative 25; Mismatches 41; Indels 37; Gaps 5;

QY 23 LIGIGILVLLITVIGVPLIFTIKANSEACRDGRAYMCRNVTLLQOELTEAQKGFQ 82  
 Db 40 VLCIGLVLVWILLLQLSQVSDLIKQQ-----ANITH--QEDILEGO--- 80

QY 83 DVEQAATCNHTWALMASLDAEKAGQKVEELEGETTTLNHLKLODASAEVERLRRNQ 142  
 Db 81 -----ILAQRSSEKS-AQESQKELEMETLAHKDESKKLMELHRONL 124

QY 143 VLSVRIADKKY--PSSQD 159  
 Db 125 NQEVLKPEANNSGCPQD 143





TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-742-923A-4

Query Match 10 7%; Score 95; DB 2; Length 885;  
 Best Local Similarity 24.5%; Pred. No. 0.11; Mismatches 27; Conservative 18; Indels 10; Gaps 1; Matches 33; Mismatches 50; Indels 23; Gaps 5;

RESULT 9  
 US-08-938-105-3  
 Sequence 3, Application US/08938105  
 Patent No. 6353151  
 GENERAL INFORMATION:  
 APPLICANT: LeJawand, Leslie A.  
 APPLICANT: Wikstrom, Karen L.  
 TITLE OF INVENTION: TRANSGENIC MODEL FOR HEART FAILURE  
 NUMBER OF SEQUENCES: 3  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sheridan Ross, P.C.  
 STREET: 1700 Lincoln St., Suite 3500  
 CITY: Denver  
 STATE: CO  
 COUNTRY: U.S.A.  
 ZIP: 80203  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08938105  
 FILING DATE:  
 TELEPHONE: (303) 863-9700  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Crook, Wannell M.  
 REGISTRATION NUMBER: 31,071  
 REFERENCE/DOCKET NUMBER: 3595-4  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (303) 863-0223  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1886 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 IMMEDIATE SOURCE:  
 LIBRARY: GenBank  
 CLONE: 1235724  
 US-08-938-105-3

Query Match 10 7%; Score 95; DB 2; Length 885;  
 Best Local Similarity 24.5%; Pred. No. 0.11; Mismatches 27; Conservative 18; Indels 10; Gaps 1; Matches 33; Mismatches 50; Indels 23; Gaps 5;

RESULT 10  
 US-08-688-342-4  
 Sequence 4, Application US/08688342  
 Patent No. 5871964  
 GENERAL INFORMATION:  
 APPLICANT: Au-Young, Janice  
 APPLICANT: Cocks, Benjamin G.  
 APPLICANT: Goll, Surya K.  
 APPLICANT: Hillman, Jennifer L.  
 TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN  
 NUMBER OF SEQUENCES: 5  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Drive  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: US  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/087688,342  
 FILING DATE: Filed Herewith  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Billings, Lucy J.  
 REGISTRATION NUMBER: 36,749  
 REFERENCE/DOCKET NUMBER: PP-0095-1 CIP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-825-0555  
 TELEFAX: 415-845-4166  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 292 amino acids  
 REFERENCE/DOCKET NUMBER: PP-0095-1 CIP  
 STRANDEDNESS: single  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 IMMEDIATE SOURCE:  
 LIBRARY: GenBank  
 CLONE: 1235724  
 US-08-688-342-4

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

RESULT 11  
 US-09-113-788-4  
 Sequence 4, Application US/09113788  
 Patent No. 5959104  
 GENERAL INFORMATION:  
 APPLICANT: Au-Young, Janice  
 US-08-742-923A-4

Query Match 10 5%; Score 93.5; DB 4; Length 1886;  
 Best Local Similarity 24.4%; Pred. No. 0.46; Mismatches 33; Conservative 29; Indels 23; Gaps 5;  
 Matches 33; Mismatches 50; Indels 23; Gaps 5;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

Query Match 10 4%; Score 92.5; DB 2; Length 292;  
 Best Local Similarity 24.6%; Pred. No. 0.043; Mismatches 32; Conservative 28; Indels 39; Gaps 4; Matches 32; Mismatches 39; Indels 31; Gaps 4;

APPLICANT: Cocks, Benjamin G.  
 APPLICANT: Goli, Surya K.  
 APPLICANT: Hillman, Jennifer L.  
 TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN  
 TITLE OF SEQUENCES: 5  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 STREET: 3114 Porter Drive  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: US  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/113,788  
 FILING DATE:  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/688,342  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Billings, Lucy J.  
 REFERENCE/DOCKET NUMBER: 36,749  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-845-0555  
 TELEFAX: 415-845-4166  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 292 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 IMMEDIATE SOURCE:  
 LIBRARY: GenBank  
 CLONE: 1235724  
 9-113-788-4  
 LT 12  
 9-111-470-4  
 quence 4, Application US/0911470  
 tent No. 6277959  
 ENERAL INFORMATION:  
 APPLICANT: Valladau, Jenny  
 APPLICANT: Ravel, Odile  
 APPLICANT: Bates, Elizabeth E. M.  
 APPLICANT: Ford, John  
 APPLICANT: Saeland, Sem  
 APPLICANT: Lebecque, Serge J. E.  
 TITLE OF INVENTION: Mammalian Membrane Protein  
 TITLE OF INVENTION: Related Reagents

Page 7

QY 59 APMECRNTHLILQQLTEAKQKFQDVQAATCNH-----TVMALMSLD-ABKA 1077  
 Db 1746 AVQECRNAAEKAKKAITDAMMAELKKQDTSAHLMKNNQETKDLQHRLDREQI 1800  
 QY 108 Q---GOKVKEELEGEITTLNKHQDASAVERURRENQVLSYRIADKKY---YPSQDS 1600  
 Db 1806 ALKGKQKQOLKLEARVLEGEI-----EAEQRNATSVKGMRKSERKILETYQERDK 1866  
 QY 161 SSAAPOLLIVLGL 175  
 Db 1861 KNLRLQDLDVQQL 1875

RESULT 14  
 US-08-312-949-4  
 ; Sequence 4, Application US/08312949  
 ; Patent No. 6027734  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Briles, David E.  
 ; APPLICANT: Wu, Hong-yan  
 ; TITLE OF INVENTION: MUCOSAL ADMINISTRATION OF  
 ; TITLE OF INVENTION: PNEUMOCOCCAL ANTIGENS  
 ; NUMBER OF SEQUENCES: 6  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Curtis, Morris & Safford, P.C.  
 ; STREET: 530 Fifth Avenue  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: United States of America  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: FLOPPY disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/312.949  
 ; FILING DATE: 30-SEP-1994  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Frommer, William S.  
 ; REGISTRATION NUMBER: 25,506  
 ; REFERENCE/DOCKET NUMBER: 454312-2049  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 840-3333  
 ; TELEFAX: (212) 840-0712  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 288 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-08-312-949-4

Query Match 9.9%; Score 88; DB 3; Length 288;  
 Best Local Similarity 30.9%; Pred. No. 0.13; Gaps  
 Matches 34; Conservative 15; Mismatches 29; Indels 32; Gaps  
 QY 60 VMECRNVTILQQLTE-----AOKGFQDVQAATCNHVTMAMASDAEKAGQG 1100  
 Db 178 IAELENQVHLQEQLEKEDESESEDYAKGFR-----APQSKLQDKAKL 2244  
 QY 111 K-----KVEELGEITTLNKHQDASAVERURRENQVLSYRIADKK 152  
 Db 225 KLEELSKDIDELAEKAKLQDQKAEEENNNDYKGG--LRTIAK 272

GENERAL INFORMATION:  
Filing Date: 004203  
APPLICANT: BRILES, David E.  
APPLICANT: WU, Hong Yin  
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS FOR MUCOSAL ADMINISTRATION OF  
FILE REFERENCE: 454312 2018  
CURRENT APPLICATION NUMBER: US/08/446,201B  
CURRENT FILING DATE: 1995-05-19  
EARLIER APPLICATION NUMBER: 08/312,949  
EARLIER FILING DATE: 1994-09-30  
EARLIER APPLICATION NUMBER: 08/246,636  
EARLIER FILING DATE: 1994-05-20  
EARLIER APPLICATION NUMBER: 08/048,896  
EARLIER FILING DATE: 1993-04-20  
EARLIER APPLICATION NUMBER: 07/835,698  
EARLIER FILING DATE: 1992-02-12  
EARLIER APPLICATION NUMBER: 07/656,773  
EARLIER FILING DATE: 1991-02-15  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 288  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-08-446-201-4

STATE: New York  
COUNTRY: United States of America  
ZIP: 10036  
  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/312,949  
FILING DATE: 30-SEP-1994  
CLASSIFICATION: 424  
  
ATTORNEY/AGENT INFORMATION:  
NAME: Flommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2049  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 288 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-312-949-4

RESULT 15  
US-08-446-201-4  
; Sequence 4, Application US/08446201B1

THIS PAGE BLANK (USPTO)